(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

POWELL MURCHISON
Managing Editor

HELEN M. DAVIS
Associate Editor

Editorial Board

GORDON ALLPORT
Harvard University

JOHN A. ANTOINETTI
Life Insurance Agency
Management Association

ARTHUR G. BILLS
University of Cincinnati

E. A. BOTT
University of Toronto

LEONARD CARMICHAEL
Smithsonian Institution

HERBERT S. CONRAD
U. S. Office of Education

JOHN F. DASHIELL
Florida State University

LEONARD W. DOOB
Yale University

ROY M. DORCUS
University of California
at Los Angeles

FRANK A. GELDARD
Princeton University

CLARENCE H. GRAHAM
Columbia University

ERNEST R. HILGARD
Stanford University

JOHN E. HORROCKS
Ohio State University

HEINRICH KLÜVER
University of Chicago

NORMAN L. MUNN
University of Adelaide,
Australia

JOHN PAUL NAFE
Florida State University

WARREN H. TEICHNER
Tufts University

LEE EDWARD TRAVIS
University of Southern
California

ROBERT C. TRYON
University of California

JOSEPH ZUBIN
Psychiatric Institute,
Columbia University

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

VOLUME 63

1966

Copyright, 1966, by The Journal Press
Provincetown, Massachusetts, U. S. A.
All rights reserved
Published Bimonthly at 35 New St., Worcester, Mass.
Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts

AUTHOR INDEX

Adams, James F. 117
Bartley, S. Howard 53, 275
Benton, Joseph G. 197
Bobbitt, Ruth A. 83
Brewer, Blayne 287
Doob, Leonard W. 13
Gengerelli, J. A. 67
Gordon, Betty N. 83
Grosser, George S. 219
Halprin, Michael 43
Hirschenfang, Samuel 197
Holman, David C. 249
Horowitz, Herbert 235
Huber, H. P. 269
Jensen, Gordon D. 83
Jordan, Nehemiah 299
Kelly, Emmet Eugene 89
Kohn, Martin 143
Kunce, Joseph T. 287
Le Furgy, William G. 73
Levenson, Edgar A. 143
Lienert, G. A. 3, 269
Luft, Joseph 261
Matarazzo, Joseph D. 249
Molde, Donald A. 249
Montgomery, Robert L. 35
Parker, Christopher E. 67
Phelan, Joseph G. 43
Pollis, Nicholas P. 35
Pyron, Bernard 159
Rowley, Vinton N. 229
Schneider, Carl W. 53
Simkins, Lawrence 201
Simmons, Mae W. 99
Stone, David R. 135
Stone, F. Beth 229
Talland, George A. 105
Teichner, Warren H. 179, 291
Walsh, Anthony A. 219
Weisberg, Paul 99
Wiens, Arthur N. 249

TABLE OF CONTENTS

Mental age regression induced by lysergic acid diethylamide 3
By G. A. Lienert

Eidetic imagery: A cross-cultural will-o'-the-wisp? 13
By Leonard W. Doob

Conformity and resistance to compliance 35
By Nicholas P. Pollis and Robert L. Montgomery

Inhibition of concept attainment in schizophrenics and in normals as a consequence of a *set* imposed by initial organization of data 43
By Joseph G. Phelan and Michael Halprin

Changes in sensory phenomena and observer criteria at low rates of intermittent photic stimulation 53
By Carl W. Schneider and S. Howard Bartley

Spectrographic analysis of electroencephalograms under conditions of alertness and relaxation 67
By J. A. Gengerelli and Christopher E. Parker

The induction of anchoring effects in absolute judgements through differential reinforcement 73
By William G. Le Furgy

Development and application of an observational method: Continuing reliability testing 83
By Ruth A. Bobbitt, Betty N. Gordon, and Gordon D. Jensen

Member personality and group counseling interaction 89
By Emmet Eugene Kelly

A modified WGTA for infants in their second year of life 99
By Paul Weisberg and Mae W. Simmons

Visual signal detection, as a function of age, input rate, and signal frequency 105
By George A. Talland

The status of psychology in the universities of Austria and Germany: 1965-1966 117
By James F. Adams

Subsumptive labeling as a variable in concept organization 135
By David R. Stone

Differences between accepted and rejected patients in a treatment project of college dropouts 143
By Martin Kohn and Edgar A. Levenson

Rejection of avant-garde art and the need for simple order 159
By Bernard Pyron

Runway performance as a function of hunger level measured as food acceptance 179
By Warren H. Teichner

Assessment of phantom limb sensation among patients with lower extremity amputation 197
By Samuel Hirschenfang and Joseph G. Benton

Effects of stimulus conditions and response contingencies on the development and maintenance of a perceptual discrimination 201
By Lawrence Simkins

Sex differences in the differential recall of taboo and neutral words 219
By George S. Grosser and Anthony A. Walsh

Children's behavior problems and mother's age 229
By F. Beth Stone and Vinton N. Rowley

The relationship of confidence to competence and imitation behavior 235
By Herbert Horowitz

Can interview interaction measures be taken from tape recordings? 249
By Arthur N. Wiens, Donald A. Molde, David C. Holman, and Joseph D. Matarazzo

On nonverbal interaction 261
By Joseph Luft

Differential effects of coffee on speed and power tests 269
By G. A. Lienert and H. P. Huber

A comparison between certain retinal and cortical activities underlying vision 275
By S. Howard Bartley

Neuropsychiatric patients, accident proneness, and interest patterns 287
By Joseph T. Kunce and Blayne Brewer

A method for studying olfaction in the unrestrained rat 291
By Warren H. Teichner

Perceived discrepancy in attitude intensity between the actors A and B in AB*X* situations and its effect upon affect 299
By Nehemiah Jordan

ERRATA

Robinson, John A. Category clustering in free recall. *J. of Psychol.*, 1966, **62**, 279-285. P. 283, the sentence beginning on line 11 should read "Reduction of the CTM in Figure 1 would yield only *18* different pairs etc." The sentence beginning on line 14 should read "Thus, the CCI for this *S* would be *.50*."

$30.00 per annum
$15.00 per volume
$7.50 per half volume

BIMONTHLY
Three volumes per year
Immediate publication

May, 1966
Volume 63, First Half

(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a liberal contract, and may not be resold for a period of five years from the date of publication.

MAY, 1966

(Manuscripts are printed in the order of final acceptance)

Mental age regression induced by lysergic acid diethylamide 3
By G. A. Lienert

Eidetic imagery: A cross-cultural will-o'-the-wisp? 13
By Leonard W. Doob

Conformity and resistance to compliance 35
By Nicholas P. Pollis and Robert L. Montgomery

Inhibition of concept attainment in schizophrenics and in normals as a consequence of a *set* imposed by initial organization of data 43
By Joseph G. Phelan and Michael Halprin

Changes in sensory phenomena and observer criteria at low rates of intermittent photic stimulation 53
By Carl W. Schneider and S. Howard Bartley

(over)

Copyright © 1966 by The Journal Press
Provincetown, Massachusetts, U. S. A.
All rights reserved

Published at 35 New St., Worcester, Mass.
Second-class postage paid at Worcester, Mass.
Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts

Spectrographic analysis of electroencephalograms under conditions of alertness and relaxation 67
By J. A. Gengerelli and Christopher E. Parker

The induction of anchoring effects in absolute judgements through differential reinforcement 73
By William G. Le Furgy

Development and application of an observational method: Continuing reliability testing 83
By Ruth A. Bobbitt, Betty N. Gordon, and Gordon D. Jensen

Member personality and group counseling interaction 89
By Emmet Eugene Kelly

A modified WGTA for infants in their second year of life 99
By Paul Weisberg and Mae W. Simmons

Visual signal detection, as a function of age, input rate, and signal frequency 105
By George A. Talland

The status of psychology in the universities of Austria and Germany: 1965-1966 117
By James F. Adams

Subsumptive labeling as a variable in concept organization 135
By David R. Stone

Differences between accepted and rejected patients in a treatment project of college dropouts 143
By Martin Kohn and Edgar A. Levenson

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 3-11.

MENTAL AGE REGRESSION INDUCED BY LYSERGIC ACID DIETHYLAMIDE*

Psychological Institute, Medical School, Düsseldorf, Germany

G. A. Lienert

A. Introduction

In two previous studies (11, 12), it was found that the profile and factorial composition of a battery of standardized tests characteristically changed under the influence of the psychomimetic drug, d-lysergic acid diethylamide (LSD: Amp. Delysid à 0.1 mg).[1] The resulting changes have been interpreted as *regressive phenomena* in the sense of age regression (*cf.* 6, p. 303).

During the interim, the sample of subjects has been considerably enlarged, enabling a more conclusive examination of earlier findings.

B. Method

1. *Subjects and Tests*

A total of 65 male and female student volunteers from the Liberal Arts and Sciences Divisions of the University of Marburg, with a mean age of 24.3 years, was given the Intelligenz-Struktur-Test (IST) by Amthauer (1). The IST consists of the following nine subtests: (*a*) General Information (SE), (*b*) Word Classification (WA), (*c*) Verbal Analogies (AN), (*d*) Verbal Relations (GE), (*e*) Memory for Concepts (ME), (*f*) Arithmetical Reasoning (RA), (*g*) Number Series (ZR), (*h*) Spatial Relations (FA), and (*i*) Cube Rotation (WÜ).

Subtest raw scores were transformed into standardized scores with a mean of 100 and a standard deviation of 10. Individual scores of the full test scale were defined by the arithmetical mean of the standardized subtest scores.

2. *Procedure*

The IST battery (parallel forms A and B) was administered to each of the 65 subjects *twice* with an interval of one week. Subjects were first

---

* Accepted by Joseph Zubin of the Editorial Board, and received in the Editorial Office, Provincetown, Massachusetts, on March 14, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] The drug was made available by Sandoz AG, Nuremburg, through the courtesy of Dr. A. Augsberger.

tested under control conditions and then under the experimental conditions involving LSD. A random sequence to the experimental condition could not be realized, since the subjects had to have prior experience with the tests; otherwise, they could not adequately perform when under the influence of LSD. Experimental conditions and positive transfer effects in the present investigation, therefore, are somewhat confounded, inasmuch as the decrease of performance under LSD is counterbalanced by the increase due to transfer. Hence, the nonrandom sequence of conditions is in favor of the null hypothesis (no change in mean level) of Prediction 1.

Each subject was tested in an individual session either in the early or the late morning.

3. *Medication*

LSD was administered orally with a cup of acidified water, two hours after breakfast. Dosage was 1.25 micrograms/kg of body weight. Tests were administered about two hours after application, as soon as the acute symptoms of the LSD intoxication (emotional reactions, hallucinations, illusions, etc.) had subsided. The testing procedure lasted approximately two hours.

4. *Design and Analysis*

The experimental conditions constituted a *three factorial design* with two fixed factors and one random factor. The first fixed factor was the treatment with the classification LSD versus control conditions. The second fixed factor was constituted by the nine IST subtests. Individual characteristics of the 65 subjects constituted the random factor.

Statistical analysis was made in the following way. First, the full-scale score of the IST was examined for changes in mean level of performance. Second, the IST subtest scores were submitted to profile analysis. Finally, scores of all nine scales were intercorrelated separately under both conditions and factored by Thurstone's centroid method (18). The resulting centroid factor matrices were then transformed to simple structure by quartimax rotation (7).

C. HYPOTHESIS

According to the *differentiation hypothesis of mental development* (2, 5), the progression of mental functioning is characterized (*a*) by increase in general level of test performance, (*b*) by alterations of test profiles in favor of abstract reasoning (*cf.* 15), (*c*) by lowering the test intercorrelations, (*d*) by differentiation of factorial composition of the respective tests, and (*e*) by reaching the simple structure of the primary mental abilities (*cf.* 17).

When LSD, as supposed, acts as a means of *regression in mental functioning,* or mental age regression, the changes should be characterized by symptoms opposite to the differentiation hypothesis.

D. PREDICTIONS

Assuming regressive developmental response to be formally equivalent to progressive development, we may derive several predictions from the combined *differentiation-regression hypothesis:*

1. The *average full-scale score* of the IST battery will be reduced under LSD, since the full-scale score is associated with progressive increments in mental development (*cf.* C-*a*).

2. The *average subtest profile* will change in level and shape under LSD so that tests of abstract reasoning will be more negatively influenced than those involving concrete operations (*cf.* C-*b*), since abstract reasoning develops later than concrete reasoning.

3. *Intercorrelations of the nine subtests* will increase under the influence of LSD, since such correlations—according to the differentiation hypothesis—decrease during adolescent development (*cf.* C-*c*).

4. *Factor composition of the subtests* will lose in degree of differentiation under the influence of LSD, since intelligence becomes more differentiated during adolescent development (*cf.* C-*d*).

With respect to changes in factorial composition, the following predictions were made through the differentiation-regression hypothesis: (*a*) the *general factor g* of the IST battery will achieve higher loadings in most subtests under the influence of LSD than under normal conditions, since *g* is assumed to decrease with age; and (*b*) the *communalities* of the subtests will increase under LSD, since they are assumed to decrease during development.

5. *Rotated factors* will lose their simple structure and become somewhat more complex (*cf.* C-*e*), since "mental organization changes from the unified general ability in childhood to a loosely organized group of mutually independent abilities in adulthood" (5, p. 373).

E. RESULTS

1. *Prediction 1*

Mean *performance level* of the 65 students was found to be 103.6 under LSD and 110.2 under normal conditions, respectively. The decrease in performance from 110.2 to 103.6 is equivalent to the decrease on the WB-I from *IQ* 123 to *IQ* 112, reported by Levine *et al.* (9) within a sample of *S*s receiving 50-200 micrograms of LSD. The difference in the present study, statistically significant at the .001 level (t test for matched samples), is equivalent to a decrease in mental age from 24.3 to 15.6 years (equivalent regression age).

2. *Prediction 2*

The ***average profile*** of the IST calculated for both conditions is given in Table 1.

TABLE 1
Mean Standard Scores ($\mu = 100$, $\sigma = 10$) of the Nine Subtests of the IST Battery Under Normal Conditions and Under LSD ($N = 65$)

	Test								
Condition	SE	WA	AN	GE	ME	RA	ZR	FA	WÜ
Normal	111.7	109.6	113.0	115.4	108.7	108.0	110.6	107.3	107.9
LSD	105.9	104.7	107.6	108.5	101.1	100.5	100.5	103.1	100.1

Profile analysis was made by the three-way analysis of variance of the individual standard scores (13). The results, presented in Table 2, may be summarized as follows:

TABLE 2
Three-Way Analysis of Variance of the IST Standard Scores with the Fixed Factors (a) Conditions (Normal *vs.* LSD), (b) Subtests SE through WÜ, and (c) the Random Factor, Subjects 1 through 65

Source of variance	Sum of squares	Degrees of freedom	Mean square	Variance ratio F
Conditions	14290.53	1	14290.53	120.21**
Subtests	9099.28	8	1137.41	12.87**
Subjects	38192.30	64	596.75	—
Cond. × Subtests	649.80	8	81.23	2.21*
Cond. × Subjects	7608.52	64	118.88	—
Subtests × Subjects	45242.61	512	88.36	—
Cond. × Subtests × Subjects	18803.65	512	36.73	—
Total	133886.69	1169	—	—

* Significant at .05 level.
** Significant at .001 level.

a. The decrease in ***profile level*** shown by the F-ratio for both conditions is equivalent to changes in the full-scale score described under Prediction 1.

b. The change of ***profile shape*** under LSD is indicated by the interaction F-ratio ***conditions*** × ***tests*** significant at the .001 level. The resulting interaction indicates that a greater loss of performance under LSD occurs, as predicted, in such tests as RA and ZR than in tests like FA, a spatial relations test. RA and ZR have been factorially identified by Fischer (4) as tests of abstract reasoning (Prediction 2).

c. To examine whether age regression in performance level and shape (Predictions 1 and 2) is consistent with age regression in performance level and profile shape, raw scores of the IST subtests achieved under LSD were transformed to standard scores according to the norms for 15-year-olds. The resulting average profile of the transformed scores did not differ statistically from the group profile under normal conditions.

d. To summarize Predictions 1 and 2, the findings are in agreement with the differentiation-regression hypothesis (*cf.* C-*a* and C-*b*).

3. *Prediction 3*

Table 3 presents the *Pearsonian intercorrelations* of the normally distributed standard scores for the two conditions.

TABLE 3
PRODUCT-MOMENT INTERCORRELATIONS OF THE IST SUBTESTS UNDER NORMAL CONDITIONS (RIGHT UPPER HALF) AND UNDER LSD (LEFT LOWER HALF)

	Test								
Test	SE	WA	AN	GE	ME	RA	ZR	FA	WÜ
SE	—	.29	.29	.22	.33	.24	.26	.35	.14
WA	.50	—	.22	.21	.10	.26	.21	.22	.03
AN	.71	.53	—	.31	.28	.45	.35	.34	.24
GE	.47	.48	.59	—	.11	.27	.15	.18	.07
ME	.24	.37	.50	.28	—	.50	.35	.35	.19
RA	.58	.43	.63	.46	.51	—	.73	.59	.45
ZR	.36	.40	.44	.34	.51	.72	—	.39	.44
FA	.37	.24	.44	.22	.34	.62	.53	—	.58
WÜ	.26	.18	.30	.17	.13	.40	.42	.64	—

a. The *average intercorrelations* raised from .30 for normal conditions to .39 for LSD. This increase was not due to an increase in standard deviations of the nine subtests, as was found by comparison of the subtest variances under LSD with those under normal conditions (*cf.* 19, p. 190). This increase of heterogeneity of performance may not have produced increase of subtest intercorrelations. The mean difference, though numerically small, may be accepted as substantial, since 30 of the 36 pairs of correlations coefficients increased under LSD conditions. The statistical sign test may not be applied for these differences, since pairs of corresponding correlations coefficients are not stochastically independent.

b. The *chronological equivalent* of the intercorrelational change cannot be estimated because differentiation theory has not yet established any function of how mean test intercorrelations change with age.

Though Prediction 3 cannot be examined statistically, numerical results provide evidence in favor of the differentiation-regression hypothesis (*cf.* C-*c*).

4. *Prediction 4*

According to Thurstone (18, p. 121) a *general factor* of a test battery may be called any factor which has substantial loadings on all (or nearly all) subtests. Two factor-analyses of the intercorrelations—one for LSD and another for normal conditions—provided evidence that the first centroid factor in each analysis could be treated as a general factor.

The *loadings of the nine subtests on the first centroid factor* are shown in Table 4.

TABLE 4
LOADING OF THE IST SUBTESTS ON THE FIRST CENTROID FACTOR UNDER NORMAL CONDITIONS AND UNDER LSD

					Test				
Condition	SE	WA	AN	GE	ME	RA	ZR	FA	WÜ
		Verbal tests					*Nonverbal tests*		
Normal	.70	.61	.80	.60	.50	.84	.73	.67	.52
LSD	.49	.36	.58	.36	.54	.83	.71	.70	.54

a. The *mean loadings* on the so-defined general factor were .67 under LSD and .57 under normal conditions. This finding is in agreement with prediction 4a and gives support to the differentiation hypothesis of intelligence factor structure as set out by the present writer elsewhere (14) (*cf.* also C-*d* above). It is interesting to note that the increase in general factor loadings under LSD was mainly due to verbal tests with low general factor loadings under normal conditions. Thus, the nonverbal general factor under normal conditions changed to a verbal-nonverbal complex factor under LSD.

b. Since the *communalities* in the factor analytic procedures are numerically dependent on the number of extracted factors—i.e., on the process of extraction and on the criterion of when to cease extraction—it was decided to compare the communalities of only the first four centroid factors under both conditions.

As may be seen from Table 5, none of the communalities under LSD is higher than the corresponding communality under normal conditions. This finding is in agreement with Prediction 4b.

TABLE 5
COMMUNALITIES OF THE IST SUBTESTS UNDER NORMAL CONDITIONS AND UNDER LSD DETERMINED FROM THE FIRST FOUR CENTROID FACTORS

	Test								
Condition	SE	WA	AN	GE	ME	RA	ZR	FA	WÜ
Normal	.40	.25	.40	.28	.40	.86	.66	.65	.56
LSD	.41	.30	.42	.29	.47	.87	.68	.66	.61

5. *Prediction 5*

Rotating the centroid matrices to simple structure according to the quartimax criterion resulted in Table 6. Results may be summarized as follows:

TABLE 6
MATRICES OF ROTATED FACTORS UNDER NORMAL CONDITIONS AND UNDER LSD

	Rotated factors					Rotated factors			
Subtest*	A_0'	B_0'	C_0'	D_0'	E_0'	A_1'	B_1'	C_1'	D_1'
	Normal conditions					*LSD*			
SE	0.02	0.17	0.18	0.38	0.46	0.13	0.25	0.72	0.27
WA	0.14	0.03	0.22	0.02	0.48	0.27	0.09	0.63	—0.11
AN	0.21	0.19	0.52	0.20	0.17	0.34	0.21	0.75	0.31
GE	0.07	0.03	0.51	0.05	0.17	0.22	0.07	0.67	—0.11
ME	0.30	0.11	0.13	0.59	0.08	0.63	0.06	0.27	0.10
RA	0.76	0.33	0.31	0.29	0.05	0.64	0.40	0.42	0.16
ZR	0.70	0.30	0.11	0.18	0.25	0.71	0.39	0.24	—0.16
FA	0.19	0.69	0.16	0.31	0.17	0.34	0.75	0.14	0.19
WÜ	0.24	0.74	0.06	0.01	0.02	0.09	0.76	0.14	—0.06
Sum of the squared loadings	1.32 = 28%	1.30 = 27%	0.77 = 16%	0.75 = 16%	0.60 = 13%	1.69 = 29%	1.57 = 27%	2.27 = 39%	0.29 = 5%

a. As may be seen from Table 6, the total variance of the tests concentrates under LSD on the first three rotated factors, whereas the variance of the tests under normal conditions is distributed almost evenly among at least five rotated factors. As a consequence, the conditions of the *simple structure* are realized more effectively under normal conditions than under LSD, a finding which is in agreement with Prediction 5.

b. Interpretation of the rotated factors results in the following conclusions: *Under normal conditions,* Factor A is a *reasoning factor* (RA, ZR), Factor B is a *spatial factor* (FA, WÜ), Factor C is a *verbal comprehension factor* (AN, GE), Factor D is a factor of *associative memory,* and Factor E may be thought of as a *verbal classification factor* (SE, WA). As predicted by Fischer (4), the IST may be analyzed for factors similar to that of primary mental ability (17). The dominating factor is that of abstract reasoning.

Under *LSD conditions,* the dominating Factor A is primarily of the *verbal type,* since all verbal tests are loaded substantially on this factor. Second to it, Factor B is of a *reasoning type* of not as clear structure as that under normal conditions. Factor C is equivalent to the Factor B under normal conditions in all respects. It is a *spatial relations* factor, not substantially influenced by the drug. Factor D defies interpretation, since there are no substantial loadings on this factor.

c. In summary, it may be concluded that under LSD the structure of rotated factors is more undifferentiated and complex than under normal conditions. Only the spatial relations factor is as well differentiated under LSD as under normal conditions. Thus, Prediction 5 tends to be verified by the procedure of factor rotation, which is in agreement with the differentiation-regression-hypothesis (*cf.* C-*e*).

F. DISCUSSION

Present findings in the intellectual area are consistent with other experimental findings in the perceptual area. Liebert *et al.* (10) have found perceptual primitivation in LSD induced psychosis, as have Krus and Wapner (8). Furthermore, the regression hypothesis is in agreement with observations of psychotherapists (3, 16) using LSD as an aid for "reliving of repressed personal memories" of early childhood. If we integrate these findings with those of the present study, it may well be that LSD is a means of inducing behavioral regression in different areas of personality to an earlier state of development. One area of this may be mental age regression.

But can it be stated with certainty whether regression is a specific effect of LSD? This question is not to be answered precisely because analogous experiments utilizing the same battery of tests produce similar results when subjects perform under alcohol and conditions involving sleep deprivation (14). In terms of this generalization it may be suggested that the regression effect is nonspecific and induced either by pharmacological or by physiological stress.

The interesting question of why regression occurs under stress may be only tentatively answered. Psychoanalysts would suggest that regression is a defense mechanism that will occur when every mental function is impaired (as under LSD) or when emotionality is aroused (as under alcohol). It may well be that successful adaptive behavior in the intellectual area is more likely to be achieved by regression than by other means of compensation.

G. SUMMARY

Sixty-five students were given a battery of verbal and nonverbal paper-and-pencil tests under normal conditions and two hours after they had

orally received approximately 0.01 mg of lysergic acid diethylamide (LSD).

Under the LSD condition (*a*) the level of performance decreased, (*b*) the shape of the battery profile changed in disfavor of the abstract reasoning tests, (*c*) the intercorrelations lowered, and (*d*) the factorial composition of the battery seemed to be less differentiated than under normal conditions. The findings were interpreted in terms of mental age regression introduced by pharmacological stress.

References

1. Amthauer, R. Der Intelligenzstrukturtest (IST) (2nd ed.). Göttingen: Hogrefe, 1955.
2. Burt, C. The differentiation of intellectual ability. *Brit. J. Educ. Psychol.,* 1954, **24**, 76-90.
3. Busch, A. K., & Johnson, W. S. LSD-25 as an aid in psychotherapy. *Dis. Nerv. Syst.,* 1950, **11**, 241-243.
4. Fischer, H. Ein Vergleich zwischem dem IST von Amthauer und dem PMA von Thurstone. *Diagnostica,* 1958, **4**, 25-32.
5. Garrett, H. E. A developmental theory of intelligence. *Amer. Psychologist.,* 1946, **1**, 372-378.
6. Hilgard, E. R. Theories of Learning (2nd ed.). New York: Appleton, 1956.
7. Kaiser, H. F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika,* 1958, **23**, 187-200.
8. Krus, D., & Wapner, S. Effects of lysergic acid diethylamide (LSD-25) on perception of part-whole relationships. *J. of Psychol.,* 1959, **48**, 87-95.
9. Levine, A., Abramson, H. A., Kaufman, M. R., & Markham, S. Lysergic acid diethylamide (LSD-25): XVI. The effect on intellectual functioning as measured by the Wechsler-Bellevue Intelligence Scale. *J. of Psychol.,* 1955, **40**, 385-395.
10. Liebert, R. S., Werner, H., & Wapner, S. Studies on the effect of lysergic acid diethylamide (LSD-25). *A.M.A. Arch. Neurol. Psychiat.,* 1958, **79**, 580-584.
11. Lienert, G. A. Pharmakologische Untersuchungen über den Abbau der geistigen Leistungsfähigkeit. Bericht über den 20, Kongress der Deutschen Gesellschaft für Psychologie in Bonn, 1955. Göttingen: Hogrefe, 1956. Pp. 144-147.
12. ———. Changes in the factor structure of intelligence tests produced by d-lysergic acid diethylamide (LSD). Proceedings of the First International Meeting of Neuropsychopharmacology in Rome. Amsterdam: Elsevier, 1960. Pp. 461-465.
13. ———. Eine varianzanalytische Methode zum Nachweis der Gruppenspezifität von Testprofilen. *Z. Exp. Angew. Psychol.,* 1963, **10**, 333-345.
14. ———. Belastung und Regression Meisenheim am Glan: Anton Hain, 1964.
15. Piaget, J. Psychologie der Intelligenz. Zürich u. Basel: Rascher, 1948.
16. Sandison, R. A. Psychological aspects of the LSD treatment of the neurosis. *J. Ment. Sci.,* 1954, **100**, 508-515.
17. Thurstone, L. L. Primary mental abilities. *Psychometric Monog.,* 1938, **1**, 1-121.
18. ———. Multiple factor analysis. Chicago: Univ. Chicago Press, 1947.
19. Walker, H., & Lev, J. Statistical Inference. New York: Holt, 1953.

Psychological Institute
Medical School
Düsseldorf, Germany

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 13-34.

EIDETIC IMAGERY: A CROSS-CULTURAL WILL-O'-THE-WISP?*

African Studies, Yale University

LEONARD W. DOOB[1]

A. INTRODUCTION

Presumably any aspect of behavior merits cross-cultural investigation to determine its incidence and correlates in different societies and hence to provide fresh insight into the moot but challenging problem of the universality and applicability of scientific principles and generalizations. Although dream content has been included in such studies, almost no systematic attention has been paid to people's normal, waking images. They have been dismissed as private, subjective phenomena whose content must be heavily cultural; nothing more, it has been assumed, need be said.

Relatively unnoticed, however, is the undocumented view, promulgated a long time ago (4, 5), that among nonliterate people a so-called primitive form of imagery, eidetic imagery, must be more prevalent than it is in the West. Although they may be associated with any sense modality, eidetic images (EI) are usually defined visually: they are images which are reported to appear in front of the eyes (whether or not the eyes are open), to persist after stimulation by an external stimulus for a period of time generally longer than an ordinary after-image, to be scannable, and to be colored positively rather than negatively (i.e., the image from a red object is red, not green). Over a half century of research in the West, especially in Germany, has shown that by and large EI are negatively correlated with age and hence discoverable

---

* Received in the Editorial Office, Provincetown, Massachusetts, on March 15, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] Deep gratitude is expressed to General Mohamed Abshir for supplying hearty encouragement, friendly advice, and transportation in Somalia; to Dr. Gordon M. Wilson of Marco Surveys Limited for intellectual and technical assistance in Kenya and Tanzania; to Anita Francke and her associates in the Nordic Tanganyika Project for permitting me to draw people and data from her own research site; to three of my four interpreters—Mohamed Forrah Jama of Somalia, Gaddiel T. Ole Salaash of Kenya, and Nasson Mulili of Kenya—for functioning also as guides and pleasant, patient, stimulating companions; to E. A. Bayne of the American Universities Field Service for inspiring the third expedition to Africa; to Professor Elsa M. Siipola for no-nonsense criticism; and to the Concilium on International Studies of Yale University and the Carnegie Corporation of New York for generously sponsoring the research. As ever, forever, I also would salute many Africans respectfully, yet joyously, though I can name them not.

among children but only very, very rarely among adults. Scattered anecdotes, however, suggest that African adults may have a "photographic memory"; thus, this investigator was once told, students at a university in West Africa allegedly reproduced, during an examination, pages from the assigned books so accurately that they hyphenated words at the end of lines in the exact manner of the original texts. The most recent and also probably the best summary of previous research in the West has been provided in German by Traxel (9), who, therefore, is cited in this article whenever necessary and possible.

EI lend themselves to cross-cultural research because, fortunately, they can be objectively investigated: unlike even simple measures assessing personality traits, those testing EI are almost as noncultural as the simple observation necessary to decide whether a person from any society has sneezed at a given moment; the grief comes when the significance attached to the act is appraised. In addition, the research apparently involves no serious interviewing problems: it is easy to induce *S*s to cooperate when they are asked merely to look at a series of drawings and photographs and report what they see afterwards. Here, nevertheless, is no research utopia devoid of monotony: each *S* must be slowly, agonizingly tested under conditions that are as dully uniform as possible.

This investigation has had three interrelated objectives besides attempting to verify or disprove anecdotes concerning African imagery:

1. *Variability:* do EI, like other processes viewed cross-culturally, vary from society to society with respect to (*a*) *incidence* and (*b*) *S*s' reports about the subjective *reality of the experience?*

2. *Functions:* if the incidence of EI in fact is higher in Africa or in some African societies than in other societies and if it can be assumed that all human beings possess the potentiality of EI, then what function is served by the ability to "see" nonexistent pictures in front of the eyes? For example, (*a*) is *recall* of details from the original situation aided? Or what (*b*) *ethnopsychological functions* are ascribed to EI; i.e., what do Africans themselves say about the phenomenon? Since no (or perhaps almost no) African society evolved a system of writing but since essential facts—such as genealogies, historical events, property rights—must be stored in any society, could EI be a partial substitute for the written records of literate societies?

3. *Correlational-causal factors:* are EI associated with specifiable demographic factors, such as (*a*) *education* and *acculturation,* or with (*b*) *psychological factors?* In addition, if a society can be located in which EI are very

frequent, can their arousal be affected by the deliberate manipulation of (*c*) *experimental variables?*

B. PROCEDURE AND SUBJECTS

Cross-cultural comparisons require at least a standardized procedure like that offered by a set of Rorschach cards. For EI, alas, investigators in the West agree only with respect to the broad outline: *S* is shown a drawing or photograph (hereafter called *picture*) for a prescribed period of time, usually between 15 seconds and one minute; after its removal, he is asked to report what he sees in front of his eyes or upon a blank screen in front of him. Otherwise, there is methodological chaos: arbitrary pictures, varying numbers of pictures, different exposure times, and fluctuating criteria for judging the reports. In the investigations reported here, the basic procedure followed one recently designed for school children in New Haven, Connecticut, by Haber and Haber (3).

The blank screen consists of a piece of neutral gray cardboard, 24 by 30 inches, which is mounted on an easel. First, *S*, who is seated about 20 inches from the screen, fixates his eyes upon a four-inch red square for 10 seconds, then he reports what he subsequently sees on the screen. Successively the procedure is repeated with a blue, a black, and a yellow square. Having thus become accustomed to seeing images—most people report negative after-images—*S* is shown a black-and-white silhouette drawing which he scans for 30 seconds; afterwards he indicates what he sees on the screen. If he reports an image, he is questioned about it; if not, he is questioned about the picture which he just perceived. Three other pictures follow: another silhouette and two in full color. The four pictures are likely to appeal to American children: the silhouettes portray a family scene and an Indian hunting (both with animals); the two in color show yet another Indian fishing and Alice looking up at the Cheshire cat in Wonderland.

Testing conditions in Africa demanded several modifications in the above procedure. Sometimes it was difficult to evoke after-images at the outset; the exposure time was then extended to 15 seconds. Or exposure to four colored squares seemed unnecessary or too time-consuming; except for the sample from the first society, the number was reduced to two. Since the Habers' pictures are culturally saturated, three photographs in African settings were used either uniformly or on occasion: a black-and-white view of African women carrying infants on their backs and metal roofing on their heads, a colored scene of an African bus and cars at a village cross-road, and another of African men engaged in wood carving. In addition, various complicated montages were infrequently employed to test detailed accuracy of EI or memory; one, for example, contains a white elephant next to a white tree and a multicolored motorcar at the top, the photograph of a smiling European woman in the center, a set of 11 random numbers immediately underneath (with backgrounds in red or black and beginning and ending with zeros), and a yellow giraffe and a green elephant next to a yellow tree at the bottom.

The investigator always sat opposite *S* at a portable table so that he could observe whether the instructions to fixate the eyes upon the colors and to scan the pictures were being obeyed and whether attention was being directed upon the screen during the reporting of EI. The interpreter was at the side of the table between *S* and the investigator. Virtually all testing occurred out-of-doors, since *S*s could not be conveniently brought to a building and since their own houses usually have no windows and lack adequate illumination; occasionally a school or other public building was used. Obviously light conditions fluctuated, but an effort was made to avoid the glaring sun. It was virtually never possible to test anyone alone: spectators from the village clustered around in a circle or, in the case of public buildings, children and sometimes adults peered through windows and otherwise made their presence felt by shouts and singing. Usually, but not always, the testing and interviewing were tape-recorded. Such mundane details are reported not only because the testing conditions may have affected the results but also because it is well to note again and again that field research requiring standardization of procedure often falls short of achieving its methodological ideal.

The absence of standardized criteria for accepting verbal reports as EI may account for the great variability in the incidence of EI cited in various studies (9, pp. 270-272); hence this methodological point requires especially careful consideration. Siipola and Hayden (8) have brought previous flounderings to a sharp focus by proposing five "primary" criteria for EI: the reporting of an image, the locating of that image "out there" (on the screen or in front of the eyes), references to positive colors, the spontaneous use of the present tense to describe the image, and "appropriate" eye movements. To satisfy two "secondary" criteria, *S*s must report that the image lasts for over 40 seconds and they must be able to give correct answers to 50 per cent of questions about the image. Among Africans it has also been found useful to ask *S* to point to the section of the blank screen on which he "sees" a detail of the original picture. Two of these criteria are "objective" in the sense that *S*s can be observed: their appropriate eye movements and pointing. For cross-cultural purposes it is necessary to be certain that the language of the *S*s makes a clear-cut distinction between present and past tenses; the African languages involved in this research make that distinction. The term *trace* will refer here to two kinds of reports: *S* sees merely a small part of the picture on the screen, or his image lasts so briefly that its appearance is noted only after its disappearance.

Haber and Haber (3) report high reliabilities in the use of their criteria by two judges working independently; similar results have been obtained here in judging samples of the African tapes. Whether or not the interview was taped, the interpreter was always asked how he classified *S*'s report during the testing situation: he rather than the investigator hears *S* directly; and one of the few advantages of having to use an interpreter is that he and the investigator can comment coincidentally in English on an *S*'s performance (provided of course that the latter does not know that language).

Finally, an *S*'s total performance has been classified in one of four ways; the criterion for each category must be variously defined, since the number of pictures in the investigations fluctuated. At one extreme is the designation of *All EI,* which means that every picture without exception elicited an EI; at the other, *No EI,* indicating that none of the pictures or, at the most, only a single one elicited EI or a trace. *Mostly EI* signifies that every picture elicited at least a trace of EI and that no more than one out of four, two out of six, or three out of seven were classified as traces rather than EI. *Some EI* means that no more than two out of four, or four out of six or seven of the pictures elicited No EI; the remainder could be either traces or a mixture of traces and EI. If no image or only a trace were reported, then *S* was questioned in the past tense about the picture itself in order to be certain that he had been attentive during its exposure.

Samples of adults were tested in five African societies on three different occasions. The initial contact was established with a group of Ibo in Eastern Nigeria ($N = 45$) because anecdotes had suggested that these people might have EI (1). Then, on the second trip, a sample from a tribe in another part of Africa, the Kamba in Central Kenya ($N = 49$), was utilized (2). The principal purpose of the third trip was, if possible, to affect EI through experimental manipulation; to do this, it was necessary to find a society where they are prevalent. Contact was established with the Somali who seemed promising because they are nomads and relatively unacculturated from the viewpoint of the West ($N = 40$); when few of them reported EI, another nomad group, the Masai in Kenya ($N = 26$), was visited; when performing the experiment there also proved impossible, yet another tribe, a mixed Swahili-speaking group in eastern Tanzania ($N = 41$), was tested, with equally disappointing results. Since time for research by then was growing short, the investigator beat a retreat to the Kamba who had previously produced a high incidence of EI; and to be doubly sure of obtaining EI, only school children were sampled ($N = 64$).

Except among the Swahili, the interpreter, always a male, was a respected member of the society and usually known to most of the *S*s; thus rapport was gained without much loss of time. With the same exception, informants were not paid. Each Masai adult, however, was offered a generous quantity of snuff immediately after testing; and the Kamba children were given candy after the entire group had been seen. All *S*s were haphazardly selected: the interpreter and the investigator approached a dwelling or village and asked for volunteers or for the cooperation of particular persons. Particular Swahili adults were sought for reasons explained in the Results section; in the interest of greater speed and the strained mental health of the investigator, they and the headmen who corralled them at a convenient place were paid a nominal sum; and then to avoid a charge of discrimination, the few haphazardly selected Swahili children were also modestly compensated.

After being tested, *S*s (except among the Ibo) were questioned about their images in general; Kamba adults were also given various performance tests. The procedure varied from society to society and from *S* to *S*, as new hypotheses were tested or as a result of the exigencies of an interview; for example, the questioning was cut short if *S* grew impatient or had to attend to his children or cattle. For this reason, the *N*s vary and are indicated in parentheses whenever necessary: the first figure represents the number responding in the indicated direction, the second the total so tested or queried.

Among Kamba children, four variables—also described in the Results section—were manipulated, each in two ways, so that *S*s were randomly assigned to 16 different experimental groups. They were recruited from three schools, within one of which older and younger *S*s were drawn; hence two schools contributed one *S* each and the third two *S*s to each of the 16 groups.

C. RESULTS

All differences called "significant" have been appraised by chi square, corrected for continuity, two-tail test, with the resulting p being .05 or less.

1. *Variability*

a. Incidence. The cross-cultural comparison of EI is summarized in Table 1 which includes as a baseline the comparable findings of Haber and Haber

(3) for children in New Haven and those of Siipola and Hayden (8) for retarded children, also American. This latter group is divided into those without ("familial") and those with brain injury according to their medical records. Among the African samples, "Children" refers to those below the age of 16; among the Kamba, "Younger children" means experimental *S*s with less than six years of education and between the ages of eight and 15, "Older children" those with seven or eight years of education and with ages ranging between 14 and 18.

TABLE 1
PERCENTAGES IN EIDETIC IMAGERY CATEGORIES

	All	Mostly	Some	No	*N*
New Haven children	8	0	47	45	(151)
Retarded children					
Familial	6	0	11	83	(18)
Brain-injured	50	0	13	37	(16)
Ibo					
Children	29	12	35	24	(17)
Adults	14	14	11	61	(28)
Total	20	13	20	47	(45)
Kamba					
Younger children	10	8	10	71	(48)
Older children	0	6	0	94	(16)
Adults	20	4	14	61	(49)
Total	13	6	11	70	(113)
Masai					
Children	0	0	50	50	(6)
Adults	5	0	75	20	(20)
Total	4	0	69	27	(26)
Somali					
Children	0	0	6	94	(16)
Adults	0	0	4	96	(24)
Total	0	0	5	95	(40)
Swahili					
Children	13	13	13	63	(8)
Adults	6	0	9	85	(33)
Total	7	2	10	81	(41)

Clearly the incidence of EI varies from group to group. Among adults, the figure for All EI is higher, but not significantly so, among the Ibo and Kamba than among the Masai and Swahili; and the incidence of some form of EI is significantly higher among the Kamba than among the Swahili and almost so among the Ibo in comparison with the Swahili. The Masai reveal a definite modal tendency to report a clear trace of EI (with all colors positive) which lasted, however, no longer than 20 seconds; in the words of one informant, the image stayed "just a little after" the removal of the picture. Except for the Somali, among whom only two instances of EI could be found, the prevalence of EI among African adults is thus not sensationally high but undoubtedly much higher than that normally reported for Westerners.

For children, Table 1 also reveals considerable variability in incidence of EI, but much of it is attributable to chance, since the size of the samples tends to be small. The incidence of All EI among the Ibo children is significantly higher than among the New Haveners. The highest percentage in that EI category, however, appears among the brain-injured children.

Comparisons between children and adults within the same African societies are probably meaningless unless educational status is held constant. The nonsignificantly higher incidence among Kamba adults than among children shown on Table 1, consequently, may reflect the significantly greater education of those children. Among the Ibo the expectation from previous research concerning the negative correlation of EI with age is confirmed only in the urban but not among the largely illiterate rural group (1, p. 360). Both age groups among the Somali were almost completely devoid of EI.

b. Reality of the experience. *S*s claiming EI reported what they "saw" in a manner virtually identical in all five societies: they gave the impression of being convinced of the reality of what they were "seeing" and, it is felt, they were not fabricating their replies to please the investigator or his interpreter. Actually *S*s were uniformly praised no matter what they said; those without EI were complimented on their "good memory" after recalling a few details; obviously they were not branded as stupid or uncooperative for not claiming EI. The evidence in behalf of the validity of the reports can be grouped under three headings:

(1). *Direct, verbal suggestion had no effect upon the incidence of EI.* Except for half the Kamba children in the experiment, all *S*s were told beforehand that they were to report what they saw on the blank screen after the removal of the pictures; if they had wished to please the investigator, they would have produced totals higher than those appearing in Table 1. Among the Somali, a deliberate effort was made to provoke EI: "Most people," some were told, "can see something on the screen after I take it away; you will probably see something too. . . ." No *S* given that suggestion, however, reported even a trace of an EI, but all could and did accurately recall many details when asked what they remembered.

(2). *Opportunities to receive additional praise by reporting EI more frequently or in greater detail were ignored.* Particularly among the Masai, there were *S*s who were praised when reporting EI at the onset of the series of pictures but who later on, in spite of the praise, stated they could see nothing. The same Masai, with a modal tendency to report very brief EI, did not apparently strive for greater praise which, they could have reasoned, would have been theirs if they had claimed EI of longer duration later in the series. Then, although asked to indicate what they saw on the screen, some Ss on their own initiative stated that the image appeared not there but somewhere in space between the screen and their eyes. Many, moreover, said that they could "see" only part of the picture; thus a 12-year-old Kamba boy noted a portion of a silhouetted man on the screen, could not "see" what the man was doing, but immediately "remembered" that detail correctly. Later in the same interview or, among Kamba adults, in subsequent interviews, very few informants reported "seeing" any images on the screen; but they also could recall the pictures and details therefrom. If they had been trying to please the investigator right after the exposure, it might have been supposed that they would continue to do so minutes or weeks later.

(3). *Objectively observable behavior seemed congruent with what would be expected as a result of "seeing" EI.* Ss followed instructions literally and tried valiantly to report something on the blank screen without necessarily claiming to have EI. Again and again the Somali stated apologetically, in the words of one member of the rural police force, "I can see nothing on the screen, but I remember the picture." Another said that he saw "specks" upon the screen, which was a reference to the roughness of the cardboard; and similarly a Masai reported "a white place and nothing more." A young Swahili male continued to stare at the screen while recalling from memory details of the pictures; when questioned, he exclaimed, "I keep looking at the screen but, I am sorry, I see nothing." Or the reverse occurred when a Masai herder, questioned about traces of EI, said, "I was surprised: there was nothing on the screen, but I could see the images." Then without exception, without previous instruction, and—may it also be said?—without ever having read psychology textbooks, all Ss with some form of EI shifted the tense of their verbs from the present to the past while reporting, immediately or subsequently, the disappearance of the images. Those not seeing EI did not follow the lead of the investigator whose inquiry was in the present tense, rather they used the past. "What do you see on the screen?" a Masai was asked in the usual way; his response began, "I saw a man. . . ." A Swahili man with clear EI: "I looked away at you, then I put my eyes back on the screen, and I found the picture was gone."

No report, furthermore, was classified as EI unless the S's eyes continued to scan the blank screen; whether or not Ss appreciated the fact that their eyes were being carefully observed is not known, but certainly their entire bodily stance remained oriented toward that screen. In addition, every S reporting EI, who was also asked to point to the part of the blank screen where the "picture" still lingered, was able to do so without hesitation or delay. There were *S*s who spontaneously noted the sudden but, more usually, the gradual disappearance of EI either while the image was fading or afterwards. A Masai said the following about his brief EI, "I saw this image on the screen, but it was moving from place to place and then disappeared. . . . It disappeared slowly like a man seeing a moon and then a cloud is coming to cover it and you are trying to show somebody else and then that cloud covers it." Or a Swahili boy of 12 while reporting an EI, "I can still see the marks of the building but it is dark . . . I cannot see the people any more." Occasionally, especially among children, *S* stared so long and conscientiously at the screen that his eyes began to water; or he would mechanically brush off a fly from his face without removing his eyes from that screen. In brief, the deep, moving sincerity with which virtually all these Africans cooperated and indeed their verbal expressions really need to be experienced to give credence to their reports, the tenor of which did not seem to vary from culture to culture.

2. *Functions*

a. Recall function. The most obvious function to be ascribed to EI is the facilitation of recall, for it has been assumed that those with EI, simply by reading off the required information from their images, could recall details of original stimuli more accurately than those without EI. In spite of the dramatic anecdotes in the literature to this effect, supporting evidence turns out to be scanty or negative. Among those with All EI in Table 1, the New Haven and the retarded children reported details and colors slightly better than those without the images (3, p. 137; 8, p. 280). If anything, Kamba adults tended to be more accurate when recalling pictures from memory than when "seeing" EI (2, pp. 9-12).

In all five African societies there were some *S*s who, though reporting EI, stated that they could not "see" in their image the detail about which they were being questioned, or else they provided an incorrect report. A single quotation from each society indicates that EI do not necessarily include every aspect of the original picture; the investigator's questions are in parentheses, factual information in brackets.

> ADULT IBO CULTIVATOR: . . . I see human beings. (How many are there, can you count them?) Yes, four [five].
>
> KAMBA WOMAN: . . . I see some people and motor cars, one man on top, with goods on top. (What is the color of the bus?) Looks like red and green [brownish and white]. (What else do you see?) Many people.

> (Do you see any trees?) Yes. (Where?) [Points correctly]. (Do you see anything else?) I see nothing more. (Can you see the license plates on the cards?) No.
>
> ADULT MASAI HERDER: . . . (Can you see their trousers?) No, I can't see them. (What do you see on the screen now?) I still can see that person, but it is not now very clear.
>
> 9-YEAR-OLD SOMALI CHILD: . . . I see the fishing boat, something at the bottom, I don't know what it is . . . (You still see it?) Yes. (What is in the water?) A big fish and three small ones. (You see the big fish?) Yes. (Is its mouth open or closed?) It is closed [incorrect]. (Do you see anything above the mouth of that fish?) No [bubbles].
>
> ADULT SWAHILI CULTIVATOR: . . . I see some numbers, an elephant, a man. (Do you see them clearly?) I see them. (Can you see the numbers?) I see the numbers clearly. (Can you read them?) I can read them. (Would you read them, please? Tell us what numbers you see on the screen now.) 0 [pause of 7 seconds] 2 [5 seconds] 2 [4 seconds] 1 [5 seconds] 7 [6 seconds] 0 again [actually: 0714653282760].

These five *S*s must be classified as having some degree of EI, but clearly their memories via the images were imperfect.

Less than five per cent of *S*s reporting EI and subsequently tested claimed to be able to see pictures on the screen when it was restored 15 minutes or later during the interview, and even then errors crept into the reports. A 13-year-old Kamba girl noted correctly various details from her EI which she "saw" after many intervening questions and tests, but she stated that "the boat is black," when in fact it was brown. Recall through EI over longer periods of time—one to three weeks—was even less frequent and also not accurate (2, pp. 11-12).

b. Ethnopsychological functions. Here information concerning the function of EI has been obtained from the *S*s themselves. Almost without exception those with EI, when questioned later in the testing session, stated that they had been amazed to see the images on the screen, that this was an experience they had never or seldom previously had (Masai 7/9; Swahili 7/8; Kamba 22/25). "I was quite surprised when the picture was there and removed and I saw another picture. . . . This has happened to me the first time, but I have seen it in films," a Swahili cultivator stated, and then added, "When I go to the cinema and then go home, the picture of the cinema comes in front of my eyes." Or one of his contemporaries who found EI upon the screen to be "something wonderful" went on to say that "it happens sometimes when I walk on the street, I can see something in front of my eyes."

When the operational definition of EI is extended, however, the question of function reappears. For a majority of the African informants, either spontaneously or when questioned, claimed that customarily they have the experience of "seeing pictures in front of my eyes," hereafter abbreviated as PE (Somali 2/6; Masai 15/21; Swahili 18/32; Kamba children 45/57). No statistically significant relation exists between the presence of EI and of PE in any of the societies.

Likewise many informants stated that often they have "pictures in my head," hereafter abbreviated as PI, pictorial images, when recalling events or people from the past (Somali 6/9; Masai 13/14; Swahili 6/16; Kamba 66/86). In addition PI were occasionally reported by the Somali to be "in my heart" and by the Swahili to be "in my mind." Again there is no relation between EI and PI, nor is there any between PE and PI. During or after the tests for EI, accuracy of recall was also not related to claims concerning PE or PI of the pictures that had been exposed. A Somali stated that he could see "the bus" in his head but, when asked to describe what "you see on top of the bus," he incorrectly replied that "nothing" was there.

From the standpoint of most informants so questioned there is a difference between remembering through the use of EI, PE, or PI on the one hand and what is ordinarily called simple memory or recall on the other (Somali 15/15; Masai 4/4; Swahili 7/8; Kamba adults 14/14). Sometimes, in fact, the contrast was drawn during the testing itself. A Kamba boy of 11 was asked about the EI he was reporting, "Do you see these numbers or are you just remembering them?"; and his reply came instantly and violently, "I see them clearly." A Swahili adult without EI, while recalling details of a picture, observed that "I cannot see them now with my eyes, but I can still remember them." In speaking generally about memory and images, *S*s made statements like the following:

> "If I went to a place where there was a dance and some people were dancing, I can remember those people but, when the picture comes, I can see them dancing as they were dancing" (Kamba girl of 17).
>
> "I can see with my eyes, but I remember with my head" (Somali nomad).
>
> "It is like a shade on the screen, it is because I do not see the real picture of the elephant but the image of the elephant is still on the screen" (Masai herder).

When asked whether "you see a triangle before your eyes when you try to think of one," 10 Kamba children were evenly divided between those claiming and not claiming to have an image of the figure.

PE, according to some *S*s who were directly questioned, appear "spontaneously" and thus, as it were, they evoke recall. According to one Swahili man, the PE "just come to my mind when I am having a rest." Further probing, however, revealed that most *S*s did not mean that images suddenly appear in consciousness and thus, as it were, compel them to recall a person or event; rather the act of remembering itself evokes and hence includes imagery. "When I try to remember something, the picture comes," a Masai herder stated—and that statement, virtually verbatim, was heard again and again in all the societies. Another Masai stated that normally he had no PE "unless I try to remember something; then they come in front of my eyes." A Kamba child: "I think and then I can see the pictures." A Swahili adolescent: "I think about the pictures, then they come into my mind." In short, some sort of internal response usually mediates the arousal of the images.

Otherwise little valid or reliable information could be elicited by requesting *S*s to compare the functions performed by PE or PI and by ordinary memory. So much depended upon the question itself. Among the Kamba adults, a small number spontaneously mentioned that PI enabled them to recall absent friends (9/41), but an additional number (29 more) agreed when the question was put to them directly. Those Kamba were also questioned closely, and they said that their PI were useful in other ways (37/38): they help remembering in general (16/38) or specifically with relation to school work (5/38), forecasting the future (6/10), and the dead (19/23). References to the dead also cropped up among the Swahili, the only other group in which the matter was raised (19/27); according to one young man, "I can sometimes see my father or my mother, or one of my relatives who have been dead a long time; they come into my mind and tell me something very important. . . . I can see them properly, my father or mother or one of my relatives, I can see him clearly in front of my eyes." Often, an *S* pointed to a function which, in the light of what others said, seemed idiosyncratic and which obviously required but could not receive deeper probing; for example, the one Masai with strong and enduring EI belonged to what the interpreter called the Wizard Clan and he stated that "it is only the pleasing things that I see in front of my eyes . . . but these things that I fear I only see them in my mind." Similarly challenging is this statement from a Swahili: "Sometimes a man who is dead long ago, I can see him before my eyes and he can tell me something, but I don't understand it."

Having EI did not serve the function of decreasing the attention required during perception in order to recall the original picture later. Without exception, whenever an informant was unable to "see" a detail in his image, including EI, or when he was told that his report was incorrect, he would offer as explanation for his failure or error the fact that his perception of the picture had been incomplete. "When I have noticed something, then I can see it later on," a Kamba child stated. Normal concentration or activity apparently was required during perception.

Finally, it can be said, the presence of PE or PI usually increased the confidence of informants in their own memories. Verbatim illustrations are given here from Masai adults, but they can be more (Kamba) or less (Somali) easily culled from the other four societies: "Sometimes I can get an image of a person; if I remember a person, I can get his image; I see him close, talking, and I can laugh"; "when I try to remember my father, it seems as if I am seeing him and yet he is dead"; "I think of my father, I can hear the advices he was giving, and when I try to look very closely I don't see anybody"; "something which I remember I can remember because I know, but something I see in my mind seems to me new and real, I can see the real thing in my mind." Likewise the investigator had the impression that an *S* reporting EI answered questions unhesitatingly and confidently, just as the Habers note that they observed "no qualification" in the speech of such an *S* (3, p. 136), and Siipola and Hayden describe their typical *S* as "attending intently to and enjoying his EI" (8, p. 280).

3. *Correlational-Causal Factors*

a. Education and acculturation. Among Kamba adults, the following demographic characteristics are *not* correlated with EI: age; sex; religion; age at which the informants, if literate, believed they had been taught to read and write; contacts with the capital city (2, p. 13). In contrast, though with an important exception, there tends to be a negative correlation between formal education and acculturation in the Western sense on the one hand and the incidence of EI on the other hand. For the Ibo one relation is quite significant: fewer of those living in town (and hence having greater contact with Western ideas and practices) had EI than those living in rural areas (1, p. 360). The negative association between EI and schooling is not quite significant either among the Ibo (1, p. 361) or the Kamba adults; and a similar finding appears among the latter with respect to EI and knowledge of English (2, p. 13). The higher incidence of EI among Kamba adults than among children, as indicated above, may result from the greater education of the younger generation in Kenya where formal schooling is now acclaimed an important value and a national asset.

The noteworthy exception involves the Swahili group. *S*s were recruited from a sample of 80 men who constituted the very strict, incredibly conscientious probability sample of another investigator and who had been previously interviewed concerning sociological and economic matters with special reference to agriculture. From the approximately 150 items recorded for each male informant, 55 could be considered arbitrarily to be related to modernization; for example, contacts with the capital city or desire to live there, rotation of crops, the saving of money, the boiling of drinking water, etc. As indicated in Table 1, only five of the 28 Swahili adults can be classified as having had at least Some EI. When a comparison is made between these five and the remaining 23 with respect to the 55 modernization items, no one of the differences by itself is significant—not surprising since the N of the EI category is so small—but a trend is evident: of the 55 comparisons, those with EI emerge 36 times with a higher proportion being "modern," with two ties ($p < .05$, sign test, though strictly speaking the items are not completely independent).

Unfortunately no uniform or reliable indices are available to rate or rank the five African societies with respect to degree of acculturation to the West. Impressionistically it might be guessed that the two nomadic groups in the study, the Somali and the Masai, have been least influenced by Western society as evidenced by the distance they live from a metropolitan area, their ways of life, the traditional clothing they tend to wear, etc. Contrary to expectation, then, it is these two groups, rather than the perhaps more heavily westernized Ibo and Kamba, which display the lowest incidence of EI.

b. Psychological processes. Conceivably the after-images ascertained in the preliminary trials might be related to EI, in fact evidence therefor exists in the literature [e.g. (3, p. 137)]. In all five African societies *S*s were classified as EI if they fell into the category of at least Some EI and as having negative after-images if they reported seeing such images after being exposed to the two colored squares (for the Ibo, who were tested four times, the criterion of three or four images was used). Among the Swahili and Kamba the relation is significantly positive, but no such relation exists in the other societies.

Only among Kamba adults was the attempt made to find other psychological correlates of EI, the results of which were completely negative. Adequate numbers were given tests of ability to abstract and to retain visual forms; inadequate numbers were tested with respect to memory for olfactory stimuli, a strange foreign language, colors, and very primitive thematic-apperception themes. In the same society, equally unsuccessful efforts were made to relate EI to the informants' own reports concerning the modalities appearing in their dreams, the storing of information from modalities other than vision, and the part of the body in which they believed their "self" to be (2, pp. 12-14).

c. Experimental variables. Sixty-four Kamba students, as already indicated, served as subjects in an experiment. Each of four variables was manipulated in two ways:

(1). *Set.* Half were given the set to report EI as a result of the normal procedure of instructing them beforehand that they would be asked, after the exposure of the colors in the preliminary series and then of the pictures, to "tell us what you see on the screen"; many of them, therefore, had the experience of seeing after-images and hence may have expected also to "see" some kind of image after perceiving the pictures. The other half were not shown the colors and hence did not report after-images; instead they were given a picture outside the regular series and were simply told to look at it; nothing was said, until after the removal of that picture, about reporting what could be seen on the blank screen. By the next picture, the first in that series, they were of course set to report something since they had been previously questioned about an image and since manipulating the second variable produced positive or negative anticipation concerning the blank screen; hence after the extra picture they were like any informant being tested.

(2). *Suggestion.* Half were told each time before the exposure of the first two pictures in the regular series that "most boys and girls your age see something on the screen after we have taken the picture away, and you probably will too"; for the last two pictures, they heard the negative suggestion that "most boys and girls see nothing on the screen." The other half received the same suggestions in the reverse order. Those without "set," as already indicated, were given no suggestion on the initial or extra picture.

(3). *Perception.* Half were asked during the 30-second period of exposure what they were actually perceiving in the picture, the other half were not.

(4). *Inquiry.* Half were subjected to the usual procedure of being questioned at some length concerning their EI or, if they had none or ones that quickly disappeared, concerning what they had seen in the picture when it was present. Only two questions were put to the other half: "What do you see on the screen?" or "What did you see in the picture?" and, after that reply, "What else do [did] you see?"

The percentages in the Kamba rows of Table 1 labelled "Younger Children" and "Older Children" represent a total of 15 students who had Some, Mostly, or All EI. Of these 15, 12 had been given the "set" to report EI and three had not (and of the nine reporting EI or a trace on the *first* picture, eight likewise had been given that "set"). "Set" as here defined, therefore, had a significant effect. No one of the other three variables by itself produced a significant difference. Two of them (the positive "suggestion" coming first rather than second; and the "inquiry" occurring rather than not occurring during the perception of the pictures) interacted positively with "set"; and the interaction of all three was also significant.

These experimental findings are rendered obscure by the fact that 11 of the 15 children categorized as having at least Some EI went to one of the three schools in the same area from which all *S*s were recruited. That school happens to be in the neighborhood of a Protestant mission, whereas one of the others is administered by Catholic missionaries, and the third is a government school. The 16 children from the school with the high EI incidence do not differ from those in the Catholic school or from the comparable group in the government school with respect to mean education or age; boys outnumber girls but not significantly so.

D. DISCUSSION

Once again actual investigation of a particular phenomenon has demonstrated the invalidity of easy generalizations about so-called primitive peoples: the incidence of EI is higher in some of the African societies than in the West, as scholars guessed it would be, but not in all of them. Variability is evident. It is noteworthy, however, that the attributes and characteristics of EI *when they do appear* do not vary from society to society and in fact are no different from those found in Europe and America. This last finding could be elaborated with respect to every single detail; thus the traces reported modally among the Masai have a counterpart among children tested in five different areas of the United States (6), and the view of Africans that EI either are mediated by memory or arrive spontaneously has been noted in Germany (9, p. 267). One tentative conclusion, consequently, is unavoidable: the ability to have EI is one which transcends culture; it may be as universal as belching or having the negative after-image of black spots in front of one's eyes after staring at the sun.

More or less current investigations of EI, including the present one, suggest that EI is a form of memory as dependent upon activity during initial perception and as prone to subsequent error as any other form of imagery or memory. The informants themselves, it has been shown, indicate that they cannot see in EI—or in PE or PI—any detail they have not deliberately perceived during the original exposure; imagery of this kind, as the wonderfully concise, precise German expression would have it (9, p. 320), is not *wahrnehmungsgemäss*. As in the West, moreover, almost all the Africans with EI did not know they possessed this proclivity as such until they were tested during the investigation. The relatively short duration of most EI—among the Kamba adults, for example, the mean was 111 seconds—and the lack of evidence for the claim that EI can be evoked "hours, days, even years" after the initial exposure (9, p. 269) argue against the possible view that the images have great adaptive value. Another conclusion, also tentative, emerges: it seems highly unlikely that EI, PE, or PI serve any important social function such as enabling people without standardized forms of writing to store the information essential for the functioning of their society.

From these two tentative conclusions, a third and even more tentative one follows: EI, and to a certain extent also PE and PI, must reflect a human ability which has survived from some earlier evolutionary state and which has become virtually functionless in modern adults. Without becoming cosmic, one can perhaps maintain that a very important characteristic of higher learning is the disregarding of details and hence the retention of only the relevant and essential. Nonverbal imagery is probably always concrete and not abstract; hence EI as now understood must be considered by and large "a liability rather than an asset" (8, p. 271). Momentarily unprovable but intriguing is the possibility that animals have this kind of imagery, even as the incidence has been shown to be high in a small group of brain-injured patients. With a few exceptions here and there, our society and human societies everywhere with their great emphasis upon verbal learning and recall have little use for nonverbal imagery.

And yet, according to a respectable, sensible philosophy of science, the fruitful assumption to make is that, if it survives, a proclivity must continue to perform a useful function. The fact that children with EI retain that capacity months later [(3, p. 137) and personal communication] indicates that EI somehow are reinforced. Out of all the data collected in Europe, America, and now Africa, one relevant generalization can be squeezed: confidence in one's memory is strengthened in the presence of nonverbal images. A picture from the past cannot be perceived when it is being recalled; but if it can be "seen" in some sense, it becomes possible to "verify" one's memory. The precise circumstances requiring an increment of confidence while recalling the past, however, are likely to vary: the person may have a damaged brain; he may be a child and feel insecure of himself, particularly in the presence of adults; as among the Kamba who attribute continuing importance to ancestors, he may be more inclined to obey the edicts of the dead when they appear in front of his eyes. Such people may not have been aware of EI until they were tested because blank screens for projecting images are ordinarily not conveniently available; but many of them were well acquainted, they said, with PE and PI whose arousal does not require an outside object.

The mediating processes for the imagery, nevertheless, certainly have not been isolated. EI are associated with cortical injury in only half of the cases studied by Siipola and Hayden. They apparently were facilitated by some unknown factor in the Protestant mission school mentioned in the Kamba experiment. They are correlated dimly or not at all with demographic factors or with other psychological processes in previous research and in this investigation. Puzzling contradictions appear, such as the tendency for EI to be negatively associated with acculturation among the Ibo and positively among the Swahili.

A variety of evidence, including objectively observable behavior, suggests that African *S*s must have been experiencing EI and were not merely responding to the suggestion of the investigator. In part this finding is supported, too, by the experiment among the Kamba in which the order of presenting positive and negative suggestions by itself produced no significant differences. That experiment, however, also offers contradictory evidence: the variable of "set" was effective, and the arousal of that "set" must be considered either a form of suggestion or else a way of ensuring that the instructions regarding the reporting of images on the blank screen have been adequately understood. It may be, consequently, that a proclivity to report EI cannot be aroused through verbal suggestion when it is weak or already extinguished but that, if still sufficiently strong, its arousal is facilitated through the exercise or the understanding resulting from the experiencing of after-images. But, then, why are those after-images correlated with EI in two of the societies and not in the other two? At any rate, two of the *S*s in the Kamba experiment provide anecdotal but pure support for the genuineness of the EI experience: they reported and satisfied the strict criteria for possessing those images without having been tested for after-imagery and without being alerted to a period of postexposure questioning beforehand.

The Kamba experiment as a whole, moreover, suggests that EI, regardless of the function they perform and of the factors in the social milieu which reinforce or extinguish them, may be affected by influences operating subtly at the time of testing. The possibility is not excluded that the variability in incidence from society to society may reflect differences in the testing techniques, such as *S*s' reactions to the interpreter, to the non-African investigator, or to the instructions. On the one hand, as already indicated, the impression is gained that *S*s were reporting as conscientiously as they could and that therefore they may have been immune to such outside influences. Thus a

Swahili *S* without EI stated later in the interview that he could not see the pictures on the blank screen, nor could he make them appear in front of his eyes or inside his head; rather "I can think of them. . . . I can see them in my mind"—and then he accurately recalled most details.

On the other hand, there were evidences of verbal confusion between imagery and remembering, probably more frequently than the interpreters revealed. For example, the Somali statement that a picture could be seen "in my heart" was, according to the interpreter, a way of using the equivalent of the historical present to say that details were being remembered; a "façon de parler," he actually called it. Some Somali claimed to have PI but, when challenged, switched to the past tense and said they were really just remembering. Adding to the confusion is the fact that there is no word for "picture" in Masai or Swahili, with the result that in those languages—and often also in Kamba—the English word is employed. A Somali nomad, after reporting no EI for the first picture in the series, was asked whether "you can see it in your heart." He answered affirmatively and supplied details, more or less correctly. To "Are you remembering this now?" he gave an affirmative reply; and then, immediately afterwards, to "And you see it in your heart?" he replied, "Yes, in my heart."

So much research on EI needs to be carried out, a large part of it cross-culturally, if the phenomenon is to be better understood. A few examples can be suggested. Since EI have been shown to be not simply biochemically impressed but to result from active perception and since, more specifically, German children have EI for some kinds of pictures and not for others (9, p. 269), the problem of motivation or interest arises. Certainly the finding that Ibo report pictures with or without African content equally well (1, p. 361) must be more carefully scrutinized; could it be that the novelty of the testing situation provided motivation to perceive any picture? Likewise feats of accurate recall from EI must be painstakingly investigated. Thus it was once noted that English children could reproduce a long German word, *Gartenwirthschaft,* which appeared in one of the pictures; the implication was that they could not possibly have memorized the word in the usual way, but that somehow it reappeared in their image from which its spelling could be read off. Decades later, a *German* child also "read" the same word from his EI, but he omitted the first *h* between *t* and *s,* which is the way the word has come to be spelled in modern German (9, p. 312); it may be surmised, therefore, either that he originally perceived and then stored the word according to his conception of its spelling or else that he committed this understandable error while reading from his EI. A few Africans with EI or PI could recall with apparently amazing accuracy two complicated montages which are tantalizingly similar in some respects and different in others; and they could answer questions about the pictures, such as "In which picture is the foot white rather than colored?" But equally impressive were the performances of some informants without either form of imagery. Finally, it seems essential to discover, in the words of an early investigator who stated that German children with EI tended in some grades but not in others to be in the top half of their class (7, pp. 159-60), whether "die Pflege dieses wichtigen Instruments der geistigen Entwicklung" can be encouraged by practice. But would any one of these studies enable us to run down the will-o'-the-wisp that EI now seem to be? This particular toiler in the vineyard of cross-cultural research is obviously baffled.

E. Summary

Haphazardly selected samples of children and adults in five African societies were tested for eidetic images in a more or less uniform manner closely resembling that used in two investigations of normal and of retarded American children. Although the incidence varied markedly, the phenomenon itself transcended culture: it was spontaneously reported in similar terms everywhere. The images neither aided recall appreciably nor reduced the attention required during the perception of the original pictures; but they and other related images seemed to increase confidence in the act of recall. No consistent or very significant relation was found between eidetic images and a variety of demographic factors or psychological processes; of four variables manipulated experimentally in one of the societies, that of set had a facilitating effect on the arousal of the images. The tentative conclusion is reached that images of this general type, being concrete rather than abstract, may represent a survival from an earlier stage in the development of man and that normally they may be but need not be activated when the individual is experiencing some special kind of difficulty in coping with the environment.

References

1. Doob, L. W. Eidetic images about the Ibo. *Ethnology,* 1964, **3**, 357-363.
2. ————. Exploring eidetic imagery among the Kamba of Central Kenya. *J. Soc. Psychol.,* 1965, **67**, 3-22.
3. Haber, R. N., & Haber, R. B. Eidetic Imagery: I. Frequency. *Percept. & Motor Skills,* 1964, **19**, 131-138.
4. Husén, T. Studier rörande de eidetiska fenomenen. *Lunds Universitets Årsskrift,* 1945, **41**, 1-128.
5. Jaensch, E. R. Die Völkerkunde und der eidetische Tatsachenkreis. *Z. Psychol. & Physiol. Sinnesorg.,* 1923, **91**, 88-111.

6. Klüver, H. Fragmentary eidetic imagery. *Psychol. Rev.,* 1930, **37**, 441-458.
7. Kroh, O. *Subjektive Anschauungsbilder bei Jugendlichen.* Göttingen: Vandenhoeck & Ruprecht, 1922.
8. Siipola, E. M., & Hayden, S. D. Exploring eidetic imagery among the retarded. *Percept. & Motor Skills,* 1965, **21**, 275-286.
9. Traxel, W. Kritische Untersuchungen zur Eidetik. *Arch. Ges. Psychol.,* 1962, **114**, 260-336.

Department of Psychology
Yale University
New Haven, Connecticut

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 35-41.

CONFORMITY AND RESISTANCE TO COMPLIANCE*

Department of Psychology, Oklahoma State University

NICHOLAS P. POLLIS[1] AND ROBERT L. MONTGOMERY

A. INTRODUCTION

There are many approaches to conformity (2), and a significant portion of them have conceptual roots in the work of Sherif (7) and Asch (1). In the approach based on Sherif's work: (*a*) the stimulus situation and the material are ambiguous, (*b*) elicited behavior centers around definition of the stimulus material, and (*c*) new learning takes place which constitutes the basis for judgment in similar subsequent stimulus situations. In the approach based on Asch's work: (*a*) the stimulus situation but not the material is ambiguous, (*b*) elicited behavior (for naive subjects) centers around resistance or compliance to social pressure, and (*c*) essentially no new learning takes place: judgment of similar stimulus material in subsequent situations is not affected. The former approach is a laboratory paradigm of psychological processes underlying norm formation in "real life" situations. The latter approach is a laboratory paradigm of "real life" situations wherein one is pressured by other people to take a stand or to express a judgment contrary to one's definition of the situation.

It is contended by the authors that it would be methodologically and theoretically profitable to use the term "conformity" to refer to those instances where individuals are behaving in keeping with previously internalized judgmental scales, and to use the term "compliance" to refer to those instances where individuals are behaving in a manner contrary to already established judgmental scales. In this sense, *socially* relevant conformity would be evidenced by behavior consistent with social attitudes which are a functional facet of various group, institutional or collective normative frameworks.

If these contentions are useful, it should be possible to differentiate conformity and compliance experimentally. In the process it should be possible to clarify the fundamental difference between the work of Sherif and Asch, and to demonstrate the need for relating the conforming process to attitudinal and group phenomena (5).

---

* Received in the Editorial Office, Provincetown, Massachusetts, on March 18, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This research was supported by NIMH Grant No. MH 10745-01, awarded to the senior author.

In this research, the autokinetic is employed as the stimulus material in a majority pressure situation similar to that devised by Asch.

As Pollis (5) has already shown that during collective interaction judgmental scales previously established under group social conditions tend to be more stable than scales previously established under togetherness social conditions (in the company of strangers), a way is suggested for implementing the testing and demonstration of the above contentions by: (*a*) using the autokinetic in two experimental sessions; (*b*) allowing for the formation of judgmental scales during Session I under differing social conditions A (alone), T (together), and G (group); and (*c*) placing subjects with scales established during Session I under different social conditions (A, T, G) in a majority pressure situation (Session II) so that their judgmental frameworks are put under compliance stress.

If these processes (conformity and compliance) can be usefully differentiated as suggested here, it is expected that compliance should be greater in the order A > T > G than in the order G > T > A, due to differential scale conformity in the order G > T > A. Differential scale conformity under the conditions of this experiment is suggested by previous findings (5) where the A-T-G distinction was related to stability of scales formed in judging ambiguous auditory stimuli.

Another index of strength of the original scale, other than the relative number of judgments still falling within the original scale (in Session II), is the actual amount of variability evidenced from judgment to judgment. Greater variability has been taken to mean greater instability of the judgmental framework with the autokinetic [e.g., Sherif and Harvey (8), Rohrer *et al.* (6), and Walter (9)], and the expectation here is that if conformity facilitates resistance to compliance pressure, then variability will be greater in the order A > T > G.

B. Method

Ninety naive subjects (freshman and sophomore students at Oklahoma State University) were employed in the experiment. Thirty subjects were assigned to each social condition on the basis of sociometric findings[2] so that: (*a*) none of the naive subjects knew any of the plants used in Session II, (*b*) none of the A and T subjects used together knew each other prior to the experiment, and (*c*) each of the G subjects went through the experiment with a fellow group member.

During Session I scales of 2 to 5 inches were established for all subjects.

---

[2] Questionnaire copies are available upon request.

This was accomplished through verbal anchoring by the experimenter prior to subjects' making any estimates—a procedure used by Pollis (5) with ambiguous auditory stimuli. A subjects established their scales alone, T subjects established their scales with a stranger, and G subjects established their scales with a fellow group member. Fifty estimates for each subject were recorded by the experimenter. As close to 24 hours later as possible, subjects returned for Session II. All naive subjects faced a planted majority of three individuals who made estimates on a planned sequence that fell within a range of from 7 to 10 inches.

In Session II, each A subject was paired with another A subject (a stranger), T subjects were paired with their original Session I partners, and G subjects were paired with their original group partners. This arrangement allowed for testing the relative capacity of A, T, and G social condition subjects to resist compliance pressure from a planted majority.

In both Sessions, naive subjects sat 15 feet from the autokinetic apparatus, and in Session II one plant always sat between the two naive subjects, with the other two plants seated in back. The order of response for all Sessions II was standardized from right to left so that the typical response pattern was: naive, plant, naive, plant, plant.

Fifty estimates were obtained in Session II for each naive subject. Plants arrived for the experiment at the same time as naive subjects, and were seated by the experimenter along with the naive subjects who were told that time pressures necessitated using more people at the same time.

Instructions for Session I were as follows:

> The results of this experiment may be useful for future space flight. So try to make your estimates as accurate as possible. When the room is completely dark, I will give you the signal "ready," and then show you a point of light. After a short time the light will start to move. It will move anywhere from a distance of 2 to 5 inches. As soon as you see it move, press the key. A few seconds later the light will disappear. Then tell me the distance it moved.

At that point, any questions were answered. Then the following parts of the instructions were reread:

> After a short time the light will start to move. It will move anywhere from a distance of 2 to 5 inches. As soon as you see it move, press the key. A few seconds later the light will disappear. Then tell me the distance it moved.

Instructions for Session II were as follows:

> When the room is completely dark, I will give you the signal "ready,"

> and then show you a point of light. After a short time, the light will start to move. As soon as you see it move, press the key. A few seconds later the light will disappear. Then tell me the distance it moved. Answer from right to left so that your estimates can be accurately recorded.

In order to collect data from naive subjects for each of social conditions A, T, and G, a total of 105 separate experimental sessions (including both Session I and II) were run.

C. Results

1. Inspection of data showed establishment of a 2-5 scale for all subjects during the last 25 trials of Session I. This constitutes a successful anchoring of all naive subjects to the desired scale as that scale was internalized under the different social conditions A, T, and G.

2. The main evidence required by the design of this experiment to support the hypothesis of differential compliance (A > T > G) can be obtained by comparing the frequency with which naive subjects maintained their original scales across the different social conditions—that is, the extent to which original scales (established in Session I) hold up in a majority pressure situation (Session II), comparing A subjects, T subjects, and G subjects.

The frequency with which a naive subject stayed within his original range under compliance pressure is defined as his stability score. That is, if a naive subject made 26 estimates falling within the 2-5 scale, he would have a stability score of 26. Since the hypothesis is that compliance is expected to be of the order A > T > G, stability score differences, in turn, should be of the order G > T > A. Mean stability values for G, T, and A are 41.20, 27.53, and 20.60, respectively. A Duncan (3) range test shows G > T to be significant beyond the .01 level, and T > A to be significant beyond the .05 level. Duncan's range test was employed because, as McGuigan (4) points out, it is most appropriate in that it permits the simultaneous comparison of all pairs at a more reasonable level of significance than all possible *t* tests considered jointly.

3. As stated earlier, it was also expected that greater variability of estimates would appear in the order A > T > G than in the reverse order. This dependent measure should be a concomitant of degree of stability of *S*'s original scale in Session II. The extent to which conformity results in stability of the judgmental process, as defined by resistance to compliance pressure and increased ability to maintain one's original reference scale, should also result in a tendency to displace each judgment at a decreased distance from the previous one. In effect, stable judgmental frameworks "unshaken" by com-

pliance pressure should result in relatively little judgmental variability under the stimulus conditions of this experiment. Accordingly, a variability score was computed for each naive subject by totaling the distance moved from estimate to estimate and dividing this sum by the total number of judgments. Mean stability values for A, T, and G are 1.66, 1.39, and 1.17, respectively. A Duncan's range test shows A > T to be significant beyond the .05 level, and T > G to be significant at the .05 level.

D. Discussion

Findings of significant consistent differential patterns of compliance lend support to the proposal that conformity and compliance should be treated as different processes. Conformity, then, can be characterized as behavior consonant with the way an individual perceives, judges, or defines a given stimulus or given social stimulus situation. Compliance can be characterized as behavior inconsistent with the perception, judgment, or definition of a given stimulus or given social stimulus situation.

Viewed in these terms, the conforming process is part of what gives the world its substantive base sociologically and psychologically. That is, the norms, values, or standards of groups, organizations, collectivities, etc., as internalized by given individuals in the form of relevant social and ego attitudes, significantly affect judgment, perception, and definition of a great range of social stimulus situations. The compliance process, on the other hand, refers to situations in which those relevant social and ego attitudes are put under stress to the point where expressed behavior contradicts what would normally be "appropriate" behavior.

The finding that standards achieved under group social conditions are more resistive to majority pressure than standards achieved under togetherness social conditions suggests that, in "real life," group-related standards or norms will undergo less displacement at the individual judgmental level than those same standards or norms achieved on the basis of togetherness interaction (attaching a constant value to the stimulus significance of a social pressure situation). This is another way of saying: (*a*) that conforming behavior defines the way individuals "really" feel about important social issues, problems and situations as mediated by given norms, values, etc.; and (*b*) that these judgmental, perceptual, and cognitive elements are relatively enduring and resistant to change.

The finding of greater variability of estimates in the order A > T > G than in the reverse order suggests that experiential uncertainty during psychophysical judgment is in part a function of the social conditions extant in the previous original learning situation. One implication raised by this finding is that the degree of experiential uncertainty should also be a part function of the nature of the original learning situation (in A, T, and G terms) where the stimulus material is psychosocial rather than psychophysical.

Finally, these obtained patterns of differential compliance provide additional evidence for acceptance of the proposal that the term "group" should be employed only in reference to an established normative and organizational social system, and not to situations in which strangers are interacting or in which a researcher has brought unrelated individuals together for purposes of an experiment.

The application of results obtained with this subject population to all possible subject populations is suggested by the following. First, while the stimulus material constitutes a universal phenomenon, it does not have any sociocultural or normative specific stimulus value. Second, experimental conditions relevant to the judgment process were such that group and collective norms specific to campus activity or American life were not activated. Only the issue of relatedness or lack of relatedness among individuals as it affects the establishment and stability of psychophysical judgment scales was investigated, under conditions of constant compliance stress.

E. Summary

The stability of judgmental scales established on an alone (A), togetherness (T), and group (G) basis in an autokinetic situation (Session I), when put under compliance pressure by a planted majority (Session II), was investigated. Differential subject conformity to original scales resulted in a significant A > T > G pattern of compliance. In addition, judgmental variability in Session II was significantly greater in the order A > T > G than in the reverse order.

References

1. Asch, S. E. Social Psychology. Englewood Cliffs, N.J.: Prentice-Hall, 1952.
2. Berg, I. A., & Bass, B. M. Conformity and Deviation. New York: Harper, 1961.
3. Duncan, D. B. Multiple range and multiple F tests. *Biometrics,* 1955, **11**, 1-42.
4. McGuigan, F. J. Experimental Psychology: A Methodological Approach. Englewood Cliffs, N.J.: Prentice-Hall, 1960.
5. Pollis, N. P. Relative stability of reference scales formed under individual, togetherness, and group situations. Doctoral dissertation, University of Oklahoma, Norman, 1964.
6. Rohrer, J., Baron, S. H., Hoffman, E. L., & Swander, D. V. The stability of autokinetic judgments. *J. Abn. & Soc. Psychol.,* 1954, **49**, 595-597.
7. Sherif, M. A study of some social factors in perception. *Arch. Psychol.,* 1935, **27**, No. 187.

8. SHERIF, M., & HARVEY, O. J. A study in ego functioning: Elimination of stable anchorages in individual and group situations. *Sociometry*, 1952, **15**, 272-305.

9. WALTER, N. A study of effects of conflicting suggestions upon judgments of the autokinetic situation. *Sociometry*, 1955, **18**, 138-146.

Department of Psychology
Oklahoma State University
Stillwater, Oklahoma 74075

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 43-52.

INHIBITION OF CONCEPT ATTAINMENT IN SCHIZOPHRENICS AND IN NORMALS AS A CONSEQUENCE OF A *SET* IMPOSED BY INITIAL ORGANIZATION OF DATA*

Department of Psychology, California State College at Los Angeles

JOSEPH G. PHELAN AND MICHAEL HALPRIN

A. INTRODUCTION

Carson (2), Hall (8), and Kew (9) adduce evidence to prove that schizophrenics *can* form the same kinds of concepts as normals. At the same time it is easy [see Mabry (10), Mattsson (11), and Richman (13)] to demonstrate many peculiarities of response. It has also been shown that schizophrenics experience difficulty in shifting from a once-acquired response to a new response [see Armitage, Brown, and Denny (1), and Crumpton (7)].

To date, the single most serious defect in studies on deficit in schizophrenia is the result of the failure to consider the different points at which breakdown may occur and affect all subsequent steps in the process. There are at least four basic points at which breakdown could occur: (*a*) in the peripheral units, (*b*) in the data-processing units, (*c*) in the central-processing units, and (*d*) in the motor responses. It is argued here that stimulation has to be received (peripherally) and subjected to initial organization (data-processing units) before it is centrally dealt with and produces a response. Our pilot study led us to the theory that the peripheral units, the central-processing units, and the effector processes are unimpaired in schizophrenia, and that the basic defect lies in the initial data processing.

A way of conceptualizing the deficit problem is to conceptualize the human operator as an information processing system. The particular aspect of this concept that concerns us is the individual's problems in initially organizing the mass of incoming stimulation for presentation to the central processing system in a meaningful, sequentially ordered pattern. Chapman and McGhie (3, 4, 5) describe the fundamental defect as relating to an inability to process incoming information efficiently. In one study they demonstrated, in processing perceptual materials, that where attention had to be selective and especially when competing sensory channels were involved, performance deteriorated.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on March 21, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

Chapman and McGhie (4) argued that the capacity for temporal integration, especially in the perception of speech, breaks down in schizophrenia. Schizophrenics are unable to screen out irrelevant incoming sensory data, and this results in an overloading of the short-term memory system (5).

Broadbent's filter theory postulates that there is a limit to the amount of information which can be processed per unit of time. While the primary processing channel is thus occupied, other information must be held in the short-term memory system. The rate at which information is processed by schizophrenics is abnormally slow. Since the short-term memory, by definition, can hold information for only a short time, the amount of stored information lost per unit of time will be much greater than that of normals. Hence, only a fragmentary part of the relevant stimulation will be successfully processed. From this it follows that, over periods of time, thought disorder or other bizarrenesses of behavior will appear.

One of the basic difficulties of schizophrenics is an inability to process relevant information that is being presented at too fast a rate to be efficiently handled, due to their abnormally slow primary rate of processing.

B. METHOD

1. *The Diagnostic Instrument*

For this investigation the authors used the most recent revision of the Holsopple-Phelan Test of Concept Attainment. Earlier forms of this instrument were used by Chard and Schwartz (6), Schwartz (14), and Phelan (12). The present test uses electronically sensitive copper plates with especially designed faces. The test, as used, has a number of advantages over the Smoke (15) and Schwartz (14) usages of sorting tasks. When 32 plates in two forms, with each varying in eight characteristics are used, the plates may be sorted into 28 combinations of four by holding any two dimensions common for a given sort.

The use of a simulated "automatic" sorting device makes the task independent of the examiner's whims. The apparatus consists of a four-compartmented, electrically wired grid with a light and a dummy electric eye in each compartment. The subject is informed by the light in the compartment when his response is correct and by the buzzer when it is incorrect.

2. *The Subject's Problem*

The task of the subject is to discover in which compartment each of a series of plates will turn on the light. The successful solution is achieved by the employment of a general principle. Of concern *only* to the examiner is the fact that the stimuli for particular kinds of error are distributed evenly and systematically through the series of plates.

The arrangement of the material makes it possible to infer from the non-verbal behavior, or from the record, what perceptual content and what internal bias may have influenced the subject in making his responses. It is possible to observe at which stage of the stimulus-organization pattern the block in concept attainment occurs.

3. *The Apparatus*

A machine was designed that consisted of four compartments, designated A, B, C, and D, from left to right. These compartments contained electrical plates upon which were mounted figured cards. The complete series of plates was numbered on the back, from 1 to 32. Twelve of these were also numbered I-1 through II-6. Both sets of numbers indicated the order in which they were to be presented and provided a cue to their construction. When the first nine of the sorting principles were to be used in the sorting experiment, the numerals 1 through 28 indicated the order of presentation of the plates for sorting. Plates 29 (A), 30 (B), 31 (C), and 32 (D) served as representative plates no matter which conceptual system was used or which sorting system was the basis for the sorting. The representative plates, with the sorting material attached and the way in which they work, were shown by the examiner to the subject. The remaining plates that were handled by the subject turned on the lights in the various compartments, depending on the particular concept that was to be formed in that particular sort. The concept was determined by the experimenter through the setting of a selector switch on the back of the machine.

4. *Material*

A wide variety of material may be used with this machine. In this experiment the authors restricted themselves to classificatory or taxonomic concepts.

C. PROCEDURE

The subject was shown the apparatus and the method of operation of the sample or representative cards. After demonstration, each of the representative cards was left on the shelf above the sorting compartment in which it belonged. The subject was told the right place for the card was where it caused the light to turn on; when any card was placed in a compartment where the buzzer sounded, the card was incorrectly placed. For example, the examiner might set the selector switch to setting Number 2, thus establishing *shape* as the correct principle. Then Plate I-1 was presented with the instructions: "You find out in which compartment this plate belongs. Try *not* to let it buzz the buzzer." After the subject had discovered that Plate I-1 turned on the light in compartment B, Plate I-1 was laid face up before compartment B and he was given Plate I-2. When he had found the compartment in which the light turned on, the preceding plate was removed and the plate he had just handled was left face up in front of its proper compartment. Thus the subject was always perceptually aware of his last choice. He should not be shown or told about earlier choices. The subject then proceeded through the series in this fashion unless he was able spontaneously to verbalize the principle on which he was working. If he did not verbalize the principle spontaneously after placing six successive cards without error, the examiner asked, "Why did you put it there?" Another card, belonging to a different compartment might be shown and the question repeated if the response was not clear. The subject was then told: "Now we will change the way it works. See if you can find out where these belong." The switch was then set to position Number 4, establishing *color* as the correct principle, and the procedure for Series II was followed as for Series I.

At the conclusion of this series, the *S* was again told that the plan of the machine would be changed. The switch was set to position Z for the establishment of the use of *numerals* as the underlying sorting principle and the Series I plates could again be used. Series II, Position W, was used for the principle of *number of figures on the plates.*

The aim was to examine *S*'s performance in a series of variations of sorting principles. In this way, characteristically individual approaches to the sorting problem could be studied. Two consecutive processes, the primitive and the conceptual levels of thinking, could be differentiated. When the plates were presented to *S*, those plates which were similar in some respect cohered and formed complexes or groups. Abstract concepts were formed when the common characteristics of the dichotomous sorts were noted and approximately designated.

1. *Subjects*

Thirty male schizophrenics at Sepulveda Veterans Administration Hospital were used as subjects. All were diagnosed as chronic, reactive, and nonparanoid. These subjects were matched with 30 adult male firemen of various Southern California communities.

Because of the nature of the pathological processes involved in schizophrenia, it is difficult to match adequately groups of schizophrenic patients and normals.

In any investigation of aspects of mental processes, intelligence presents itself as the factor that should be equated for both groups. Among normal groups, this can be done by matching on the basis of *IQ*. A measured *IQ*, however, does not mean the same thing among psychotics as it does among normals. It is therefore impossible to measure on this basis with the use of present techniques. If total *IQ* is used as a measure of present functioning level, excessive interest variability of the schizophrenics would make individuals in the psychotic group favored in some areas and penalized in others. It seemed most germane to the purposes of this particular investigation to restrict matching devices to two subtests of the *Wechsler Bellevue Intelligence Scale* which seemed most pertinent to the materials used in this investigation—the vocabulary and arithmetic reasoning subtests. These were employed as a basis of matching for intellectual functioning.

Wechsler (16) found vocabulary to have the highest correlation with total *IQ* of any of the subtests, and other investigators have found that the ability to attain a certain score on vocabulary tests is the least impaired of all examined abilities of schizophrenics. They have found that, when mental age is held constant, the schizophrenic is favored over the normal in taking a vocabulary test. Averaging the weighted scores of the vocabulary and the arithmetic reasoning subtests offset this advantage in favor of the schizophrenics, as Wechsler (16) indicated that there is some impairment of arithmetic reasoning among schizophrenics.

In our distribution of the two experimental groups on vocabulary and on arithmetic subtest scores, none of the differences among the groups is significant. Groups were also matched for age, educational, and socioeconomic status.

The mean age of the normal subjects was 42.7 years (their ages ranged from 32 to 58 years). The mean age of the schizophrenic subjects was 44.3 years (their ages ranged from 29 to 65). A *t* test indicates this difference is not significant. There was no reliable difference between the normals and the schizophrenics on the intelligence estimates derived from the vocabulary and the arithmetic subtests of the Wechsler Scale. The normals had a mean standard score of 11.17 with a range from 7 to 14. The schizophrenics had a mean standard score of 10.81 with a range from 7 to 17.

In order to determine whether the tests used in the present investigation were capable of providing the necessary and sufficient conditions for patients to attain concepts, a preliminary study was undertaken. Throughout the literature, the main measuring instrument for the investigation of concept formation has been the sorting test.

Unaccounted for in the previous procedural designs is the ability to retest concept formation on equated forms. The Revised Holsopple-Phelan Test presents designs which can be sorted, using different principles (large and small forms, color of the large numbers, position of the figures, background, and edge details), thus affording insights into the concept formation process. For all these purposes, a concept formation testing schema must demonstrate the employment of interchangeable concept formation tasks, standardized as to levels of difficulty.

Ultimately, it is our hope to make available to clinical psychologists, for diagnostic testing, a concept formation testing instrument that will add interchangeability and flexibility of administration to all aspects of previously utilized designs.

2. *Control of Practice and Interference Effects*

Effects of equivalence of the concept sorting tasks, and of practice and interference, could be of special importance in the seven different plate-sorting tasks from which the concept attainment abilities of chronic schizophrenics were measured. Accordingly, in order to equate practice effects and equivalence of task in this series of concept attainment tasks, each experimental group was divided into two smaller groups and the sorting tasks were rotated among them. Half of each experimental group was presented problems which first utilized *position of the numbers* and the *shape of the numerals' background;* the other half of each experimental group was presented with tasks which first involved the *shape of the large figure,* and the *color and brightness of large and small figures.* Intersort reliabilities were calculated. Performance on any one concept attainment task was not significantly different from any other, indicating that the difficulty level of the several concept formation tasks was approximately the same. At the time, the tasks were presented in reverse orders. These scores were pooled within each group to cancel out the effect of practice.

The difficulty that next presented itself was that any approach to concept attainment that was established by one series might become a source of interference with optimum performance on those concept attainment tasks that followed. To avoid these interferences, tasks were introduced in random order to break sets or approaches. These tasks are termed "practice tasks" in this experiment because they were utilized at the beginning to acquaint the subject with the nature of the problem.

In order to determine whether the experimental procedure as outlined above was feasible, and to determine the efficacy of the tests that had been adapted and constructed for this study, as well as to obtain a tentative appraisal of the validity of the hypotheses tested, a preliminary study was conducted. For this study, 15 schizophrenic patients from a chronic ward in the Sepulveda Veterans Administration Neuropsychiatric Hospital and 15 nonhospitalized normal adult subjects were utilized. These two groups were matched on the three factors mentioned above: age, educational level, and combined vocabulary-arithmetic reasoning scores.

As a result of this study it was found that (*a*) on the seven tests for concept attainment (on the basis of one-attribute principle) there occurred a to-be-expected spread in approach to the concept formation problem, and (*b*) the range of primitive to adequate abstract concept attainment did not differ significantly from one sample population to the other. This series was concluded to be adequate for testing the hypotheses in question.

3. *Validity and Reliability of Tests Used*

The preliminary study referred to above indicated that the tests used in the present study were capable of providing the appropriate stimuli for concept attainment among the subjects in the experimental groups. Since these were tests for the attainment of concepts, the range and success of actual attainment was, in fact, a measure of the validity of the tests. In this case the best manner of obtaining the coefficient of reliability was by test-retest, using the same plates but demanding different, equivalent solution by any one of the seven alternate sorting principles. Techniques involving split-half of other similar methods were deemed not advisable because the items in these tasks were not sufficiently independent of one another.

Retest situations, since they employ a different principle for sorting, are different from a first testing. The same spread or range of adequacy of task solutions should be expected from a test-retest reliability measure when a different sorting principle is utilized on each retest. Whether schizophrenics, regardless of variability within individuals, are likely to employ adequate abstract principles in concept attainment tasks in varying situations can be determined by the use of seven different tasks employed for the attainment of such concepts.

D. RESULTS

Basing their conclusions on systematic observation of *S*s as they went from one sorting task to another sorting task, and on an analysis of their time, errors, and relative success in sorting, the authors affirmed that older normal *S*s were challenged by the complexity of the stimuli presented. Such *S*s used the feedback of buzzer signals to separate out the irrelevant cues and to restate and to re-emphasize the available redundant relevant cues. In concept attainment, redundancy of cues within the same task situation facilitates the task. The increased stimulus variability of this particular task makes it more difficult than the Wisconsin Card Sorting Task. Salience of stimulus variation within any one dimension is a factor in speed and accuracy of concept attainment. In verbalizations, *S*s did not report more than one dimension in the sorting. Apparently *S*s focused on one dimension only at any one time.

There occurred some probability matching of the cards. But *S*s groped for the relevant cues in a search for the underlying principle. The trend was toward an all-or-none stratagem. The *S* set up hypotheses that were either correct all the time or incorrect all the time, or sometimes correct and sometimes not. In each presented instance *S* selected one hypothesis which was rejected if it was found to be wrong. The probability of an error at a given trial turned out to be a decreasing function, dependent upon the proportion of hypotheses that were sometimes correct and sometimes not.

In the first trials, errors occurred because of sampling a strategy which turned out always to be wrong. The strategies were systematically eliminated after the first few trials. The normal *S* tended to approach the status of an optimal performer, quick to recognize which hypothesis was being tested and when it was correct. *S*s appeared to eliminate once and forever each hypothesis when it was found to be incorrect.

Schizophrenics responded to the same task in qualitatively different ways. The research in psychological deficit pointed to the paranoid subtypes of schizophrenics as being "overinclusive" in approach in scanning and given to field articulation types of response, while minimal scanning and undifferentiated field articulation responses characterize the chronic, reactive, nonparanoid types. In the authors' observation of the 30 patients, when engaged in the alternative, complex sorting tasks, the schizophrenics were overwhelmed with stimulation which they could not handle. They appeared to screen out stimulation protectively. When they hit upon one way of sorting that seemed error free, they developed a *set;* they clung to one method of sorting all cards the same way in spite of having been told the first principle of sorting was no longer applicable.

The initial principle of organization of the first stimuli presented was adhered to rigidly throughout. Of the 30 chronic schizophrenic subjects, four were able to shift *set* in a nonreversal way and to explore systematically three distinct principles of sorting. Eighteen of the 30 were able to organize initial stimuli to the point that they developed at least a minimally adequate concept for the first sort; thereafter they maintained this as a rigid *set* and refused to shift to other principles in the face of repeated failures. Two *S*s tried to move to a nonreversal shift, but only feebly, taking half-hearted stabs at moving from their original strategy, but reverting back to their first hypothesis. Six *S*s were either unable to comprehend the nature of the task or went about it in utterly inappropriate ways. Schizophrenic impairment was noted in irrelevant and incoherent responses, and in marked dissociation with fantastic elaboration of ideas. The difference in efficiency of concept attainment between the subject populations was significant.

E. CONCLUSIONS

1. The Holsopple-Phelan Concept Attainment task can be seen as a clinical research tool that is able to provide insights into the concept attainment process in normal subjects and in psychotics. Observed patterns point to the development of strategies in the concept attainment process.

2. Chronic, reactive, nonparanoids exhibit a slowness in organizing initial data, a marked tendency to screen out stimulation protectively, and a rigidity of *set* in clinging to the principle or rule that they achieve, applying it to subsequent tasks persistently in the face of repeated failure.

3. Normal *S*s behave as at least adequate performers in concept attainment in a structured, nonverbal task, in which parameters have been introduced specifying the ultimate rejection of an incorrect hypothesis. Learning is of the all-or-none strategy type, comprising the systematic acceptance or rejection of hypotheses as they test out correct or incorrect.

REFERENCES

1. ARMITAGE, S. G., BROWN, C. R., & DENNY, M. R. Stereotyping of response in schizophrenia. *J. Clin. Psychol.,* 1964, **20**, 225-230.
2. CARSON, R. C. Proverb interpretation in acutely schizophrenic patients. *J. Nerv. & Ment. Dis.,* 1962, **135**, 556-564.
3. CHAPMAN, J., & McGHIE, A. A comparative study of disordered attention in schizophrenia. *J. Ment. Sci.,* 1962, **108**, 487-500.
4. ————. An approach to the psychotherapy of cognitive disfunction. *Brit. J. Med. Psychol.,* 1963, **36**, 253-260.
5. ————. Echopraxia in schizophrenia. *Brit. J. Psychol.,* 1964, **110**, 365-374.
6. CHARD, R. D., & SCHWARTZ, B. D. A new modified concept formation test. *Amer. Psychol.,* 1947, **2**, 274-275.
7. CRUMPTON, E. Persistence of maladaptive responses in schizophrenia. *J. Abn. & Soc. Psychol.,* 1963, **66**, 615-618.
8. HALL, G. C. Concept attainment in schizophrenics and nonpsychotics as a function of task structure. *J. of Psychol.,* 1962, **53**, 3-13.
9. KEW, J. K. A comparison of thought processes in various nosological groups. *J. Clin. Psychol.,* 1963, **19**, 162-166.
10. MABRY, M. Language characteristics of scattered and non-scattered schizophrenics compared with normals. *J. of Psychol.,* 1964, **57**, 29-40.

11. Mattsson, P. O. The Stein sentence construction test and cognitive disturbance: A cross validation study. *J. Clin. Psychol.*, 1964, **20**, 368-369.

12. Phelan, J. G. Effects of attempts to verbalize on the process of concept attainment. *J. of Psychol.*, 1965, **59**, 283-293.

13. Richman, J. Symbolic distortion in the vocabulary definitions of schizophrenics. *J. Genet. Psychol.*, 1964, **71**, 1-8.

14. Schwartz, B. D. Certain relations between verbalization and concept formation. Unpublished Doctor's thesis, Princeton University, New Jersey, 1948.

15. Smoke, K. L. An objective study of concept formation. *Psychol. Monog.*, 1932, **52**, No. 4.

16. Wechsler, D. Measurement of Adult Intelligence (3rd ed.). Baltimore: Williams & Wilkins, 1944.

Department of Psychology
California State College
Los Angeles 32, California

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 53-66.

CHANGES IN SENSORY PHENOMENA AND OBSERVER CRITERIA AT LOW RATES OF INTERMITTENT PHOTIC STIMULATION*[1]

Department of Psychology, Michigan State University

Carl W. Schneider and S. Howard Bartley

A. Introduction

Intermittent (repeated) stimulation has proved to be important and useful in studying certain characteristics of sensory response mechanisms. However, certain misunderstandings have arisen in its use and in the interpretation of results. The purpose of this paper is to discuss the phenomenology produced by low rates of photic intermittency and to relate this phenomenology to various interpretations made about brightness findings.

Intermittent photic stimulation produces several distinct types of sensory phenomena. These types appear at various frequency ranges. The obvious point at which to distinguish between single photic pulses and trains of photic pulses (intermittent stimulation) is the temporal separation between members of a train, which is just long enough to produce a response to each pulse that is quite like the response to a totally isolated single pulse. The ranges are (*a*) one in which distinct light and dark target fields alternate, the prominence of either depending upon the relative duration of the photic pulses and the null periods between; (*b*) a range in which the two phases are not fully distinct but create a sequence in which the fluctuation between light and dark is great; (*c*) a range producing flicker superimposed upon a steady brightness and thus a result in which the observer is fairly able to compare this "average" brightness with the brightness of an adjacent steady target; and (*d*) a range throughout which the fluctuations are absent and the field presents a steady brightness. This last result is often called fusion.

There is nothing extremely subtile about this four-range classification, nevertheless certain operational (methodological) confusions have been persistent over the past three decades. This has led, in some cases, to misunderstandings and to misinterpretations of findings.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on March 28, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported by Public Health Service Fellowship 5 F1-MH-17,006-02 from the National Institute of Mental Health and by Public Health Service Research Grant NB 05260-02 from the National Institute of Neurological Diseases and Blindness.

The phenomenon which has seemed to give rise to the most misunderstandings is what is called *brightness enhancement*. Bartley (4), possibly the first to use the term, defined it as the increase in brightness, produced by repetitive photic stimulation, that is above the brightness produced by steady stimulation of the same luminance.

One of the examples of confusion has been the identification of the Broca-Sulzer effect with brightness enhancement. The Broca-Sulzer effect (18, 19) is the increase in brightness manifested by a single flash as the duration of the stimulus pulse is lengthened. Although the two cases involve something that one can call an *increase* in brightness, they are by definition two separate phenomena. The use of brightness enhancement as the label for the Broca-Sulzer effect ignores, first, the fact that a technical delimitation was made in the original definition and, second, the inappropriateness of talking about enhancement without a rational reference point from which to start. A short pulse produces a given brightness; a longer pulse produces a brighter flash. If the term enhancement is proper here, it is proper in any case one chooses and, thus, the term loses the specific meaning which it was originally given. The authors contend that there is still good reason for specifically labelling the result produced by a specific set of conditions, even though some generic term may be needed finally to combine cases.

Boynton (17) has suggested that it is in error to view the Broca-Sulzer effect as an enhancement phenomenon, and recently Katz (23) ruled out Bartley's description of the processes underlying the brightness enhancement phenomenon as an explanation for the Broca-Sulzer effect.

We shall bypass, for the moment, the question of whether the same mechanism was demonstrated for producing the two cases of alleged brightness enhancement. Such has been the allegation, and even there some misunderstanding exists. The foregoing is but one example of failing to take into account the significance of distinguishing between the phenomenology of sensory phenomenon—both in experimentation and in interpretation of results.

The other prominent type of misunderstanding relevant here is exemplified by those instances in which various workers have differed in the stimulus rate at which they found maximum brightness to result. These differences, in and of themselves, would be no reason for notice here, except that such findings are erroneously used as evidence against a theory that was developed to account for enhancement as it was originally defined.

One origin of the present concern dates back to the earlier work of Bartley. In the 1930's, while virtually all study of intermittent photic stimulation was directed toward the ascertainment of CFF under various combinations of conditions, Bartley became interested in the possible results of using input rates below CFF (subfusional frequencies). His first concern was with the possible qualitative sensory effects produced.

In 1939, Bartley reported on various visual phenomena produced by intermittent stimulation at rates below fusion (CFF). Although the paper reported his study in some detail, the work was never followed up in the same vein by him or by anyone else. Later, when he reported on the increased brightness of a target resulting from intermittent stimulation, he was reporting on the brightness produced at rates below fusion which still provided a major steady brightness component despite some superimposed flicker (range *c* above). He avoided rates so low that they would result in alternation of light and dark phases extreme enough to constitute separate items, whose individual brightness would then be reported by observers.

The finding of brightness enhancement as a definite and quantifiable phenomenon sparked a long series of studies which continue to the present time. The sensory phenomenon of brightness enhancement, once it was quantified, was seen by Bartley to be potentially relatable to the neurophysiological findings of Bishop and Bartley in their study of the neurophysiology of the optic pathway in the rabbit and other species (3, 5, 9, 10, 15). The neurophysiological findings formed the basic material for what has been called the *alternation of response theory*. This theory pertains not only to the activity of the pathway, but constitutes a two-fold description of neurophysiology and sensory-end result.

In the alternation of response theory (7, 8, 11), it is postulated that peak enhancement will occur at certain critical input frequencies which are synchronous with the neural periodicity of the organism. In man, this is presumed to be about 8 to 12 cps. While some investigators (2, 4, 12, 13, 22) have obtained brightness enhancement at the critical range suggested by the theory, others have found frequencies from 2 to 5 cps to be more effective (1, 6, 12, 14, 16, 20, 21, 24, 25, 28) in producing brightness. These findings suggested to some investigators that brightness enhancement may occur anywhere between 1 to 12 cps: i.e., at rates conflicting with the theory. Such an interpretation does not recognize the fact of the phenomenological ranges described above. As a consequence, such investigators failed to realize that the brightness results obtained in the several distinct ranges cannot be compared in the manner that they attempted. It is to be expected that the brightness levels obtained with low frequency inputs are associated with single pulse phenomena, rather than with the enhancement phenomenon arising when separate flashes are not produced by the rates used.

The confusion in regard to brightness enhancement seems to be assignable to two origins: (*a*) the investigator in his conceptualization of the matter, and (*b*) the observer in the experiments themselves. While the authors have alluded to the first origin, something more needs to be said about the latter.

At the lower rates of input, the observer is unable to maintain the criterion used at the higher rates. This becomes quite plausible when one carefully considers the criterion that an observer uses when making matches under conditions of intermittent stimulation, and how that criterion may affect obtained brightness levels if it is altered. Anyone who has attempted to make brightness matches under these conditions knows that with fast frequencies the match to equality is made to some sort of average between the light and dark phase of the cycle. At the low frequencies, the observer finds it impossible to separate the two phases if called upon to do so, while the task of viewing the field as a relatively stable area is a relatively simple one. However, as the frequency is decreased, a point is reached at which the observer sees discrete light and dark phases (26) and may find it necessary to make matches to the light phase rather than to an "average" between the two: i.e., to the steady component seen. Thus, we have what may be called a *shift in criterion*. Colgan (20) in one of his studies showed that if an observer is instructed to match to the "on" phase of a pulse, there will be a significant increase in the obtained brightness levels. Similarly, we would expect that a shift in criterion, from matching to an "average" between the light and dark phase (the steady component) to matching only to the light phase, will result in a large increase in the obtained brightness level; and it is at this point that the observer is performing operations normally associated with single pulse phenomena rather than with flicker. At least, there is no basis for using the increased brightness here to contradict the interpretation of the brightness results obtained at the increased rates (10 per second, for example).

Various investigators both in vision and audition have noted changes in criteria at low rates of intermittent stimulation. Taubman (30, 31), in attempting to deal with auditory and visual number, noted that there was a change in the observer's criterion of judgment at low rates, in such a way that *time* instead of *rate* was being judged.

Reese (27), in attempting to construct an equal unit scale for visual rate, found a discontinuity in judgments when the rates between 5 and 7 per second were involved. He, too, suggested that his observers were judging "time between pulses" at the slow rates, and "rate" at input rates above 7 per second. While the tasks in these two investigations were not quite the same as we are discussing here (the judgment of brightness), they do give evidence that low rates of intermittency are reacted to in a different manner from high rates.

All of this is in line with the four-fold classification of sensory effects the authors have already made, and it suggests that the matter of shifts in sensory phenomena are involved in the contradictory experiments on the so-called enhancement maximum. What the authors are asserting here is that not all of the investigations cited, in which highly diverse rates for a brightness maximum were found, are investigations of brightness enhancement as it was originally defined; for as soon as the observer must use one phase rather than some blend or average of the two phases in the intermittency cycle as a criterion, the task shifts from the one originally defined and becomes more nearly the task of dealing with a series of separate inputs which are to be treated more nearly as *single* isolated inputs. In this respect, the task approaches the task of the observer in the original Broca-Sulzer experiment, and has definitely departed from being the task of judging brightness under conditions relevant to the alternation of response theory.

The authors have some additional evidence to offer at this time, regarding the disjunctiveness between the sensory phenomena produced as rate is varied at the lower ranges of intermittency.

B. APPARATUS AND PROCEDURE

The apparatus is shown diagrammatically in Figure 1.

The photic sources (S_1 and S_2) are Sylvania glow modulator tubes. S_1 is the steady source and is driven electronically by a modified Grass S-4 stimulator. S_2 is the intermittent source and is driven electronically by a Grason-Stadler flicker-fusion apparatus. W_1 is a calibrated neutral density wedge which could be moved to the right or left by a rack and pinion device, by turning the dial (D). W_2 is a stationary balance wedge. All wedge readings obtained from an observer when making a match were converted to percentage of transmission required for a match. This device was accurate to within .1 per cent. RS_1 is a reduction screen, and M is a first surface mirror which cut S_1 in half and reflected half of S_2. F_1 held a permanently fixed .6 neutral density filter which limited the luminance of the intermittent source to a maximum of 440 c/ft^2. L_1 and L_2 are collimating lenses which were adjustable vertically and in the line of regard. RS_2 is a field stop which was used to control the target size, and F^2 held the neutral density filters which enabled the experimenter to shift intensity levels during the investigation. RS_3 is a field stop for reducing stray illumination. Part B of the figure represents the appearance of the bipartite field as seen by the observer. FS is the intermittent

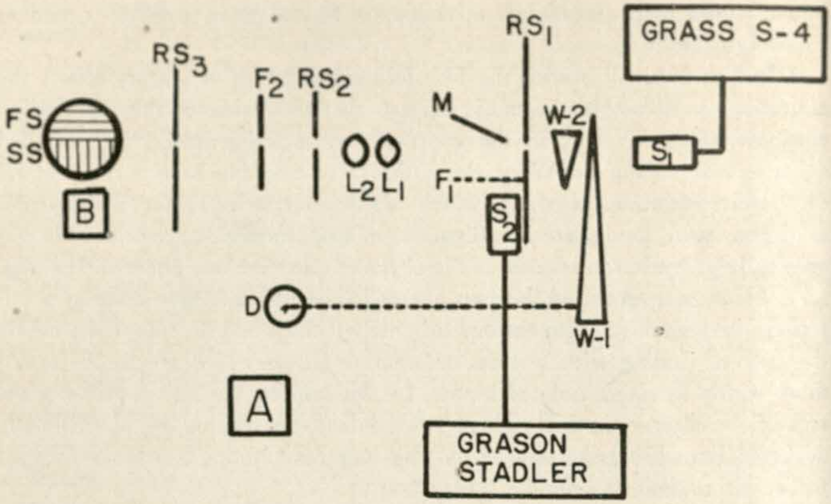


FIGURE 1
SCHEMATIC DIAGRAM OF THE APPARATUS
(For details see text)

half of the target and SS is the steady half. Under all conditions, the observer matched the steady half of the target to the intermittent half by manipulation of the neutral density wedge.

Brightness matches were made at frequencies of 20, 10, 6, 5, 4, 3, 2, and 1 cps. The pulse length, target size, and luminance were held constant at 20 milliseconds, 3.5° and 130 c/ft^2, respectively. The observers were two male graduate-psychology students who had extended previous experience in brightness-matching studies.

Each observer made 40 observations under each experimental condition in four separate 45-minute sessions. In each session, the observer made five ascending and five descending matches under each frequency condition. The conditions were presented in random order. The observer also made 10 presession and 10 postsession matches when both halves of the target were steady. At least 10 minutes were allowed for adaptation to the target luminance prior to the session.

All values obtained with the neutral density wedge were converted to percentage of transmittance of energy required for a match. The mean of the value of the 40 observations made under a particular condition was used as the percentage of transmittance required to make the two targets appear equally bright. The mean percentage of transmittance required for a match when both targets were steady was based on 320 observations, since this experiment was part of a larger investigation (29).

Expected percentages of transmittance required for a match were computed by extrapolation from the data according to Talbot's law. This operation was performed by multiplying the pulse-to-cycle fraction (PCF) by the mean percentage of transmittance required for a match when both targets were steady. Since the pulse duration was held constant in this experiment, the PCF changed with each frequency, and this was taken into account in determining the expected levels. The obtained mean percentage of transmittance required for a match at each frequency was then divided by the expected percentage to obtain what we have referred to as a ratio of effectiveness, which is only an ordinal indicator of how much more or how much less effective one condition was than another in producing brightness.

C. RESULTS

The mean ratios of effectiveness for the two observers at each frequency are shown in Figure 2. In general, there is a gradual increase in the ratio for both observers to about 6 cps, at which there is some leveling off until 3 cps is reached. At 2 and 3 cps there is a very large increase in the ratio for both observers.

The standard deviation (*SD*) of 40 observations for both observers at each frequency is shown in Figure 3. At 20 cps the *SD* is relatively small but reaches a maximum at 5 cps for both observers. By the time 2 cps is reached, the *SD* has decreased considerably but does not drop to the level obtained at 20 cps.

It was shown in Figure 2 that the ratio of effectiveness tends to level off somewhat between 6 and 3 cps for both observers. This was the range of frequencies characterized by an increase in the *SD*. In Figure 4 the means of each session are plotted for observer JK, and these tend to form a bimodal distribution, particularly in the frequency range where the ratio of effectiveness levels off and the *SD* is largest.

These findings suggest that the observer is shifting his criterion. Between 6 and 3 cps there is a tendency to use one criterion during some sessions and a different one during others. At 2 and 1 cps the shift in criterion is less evident as indicated by a decrease in the asymmetry of the distribution and the decrease in the *SD*. The *SD* is still larger at these frequencies than at 20 or 10 cps, but a great deal of this increase in variability is attributed to the fact that the task is increasingly difficult.

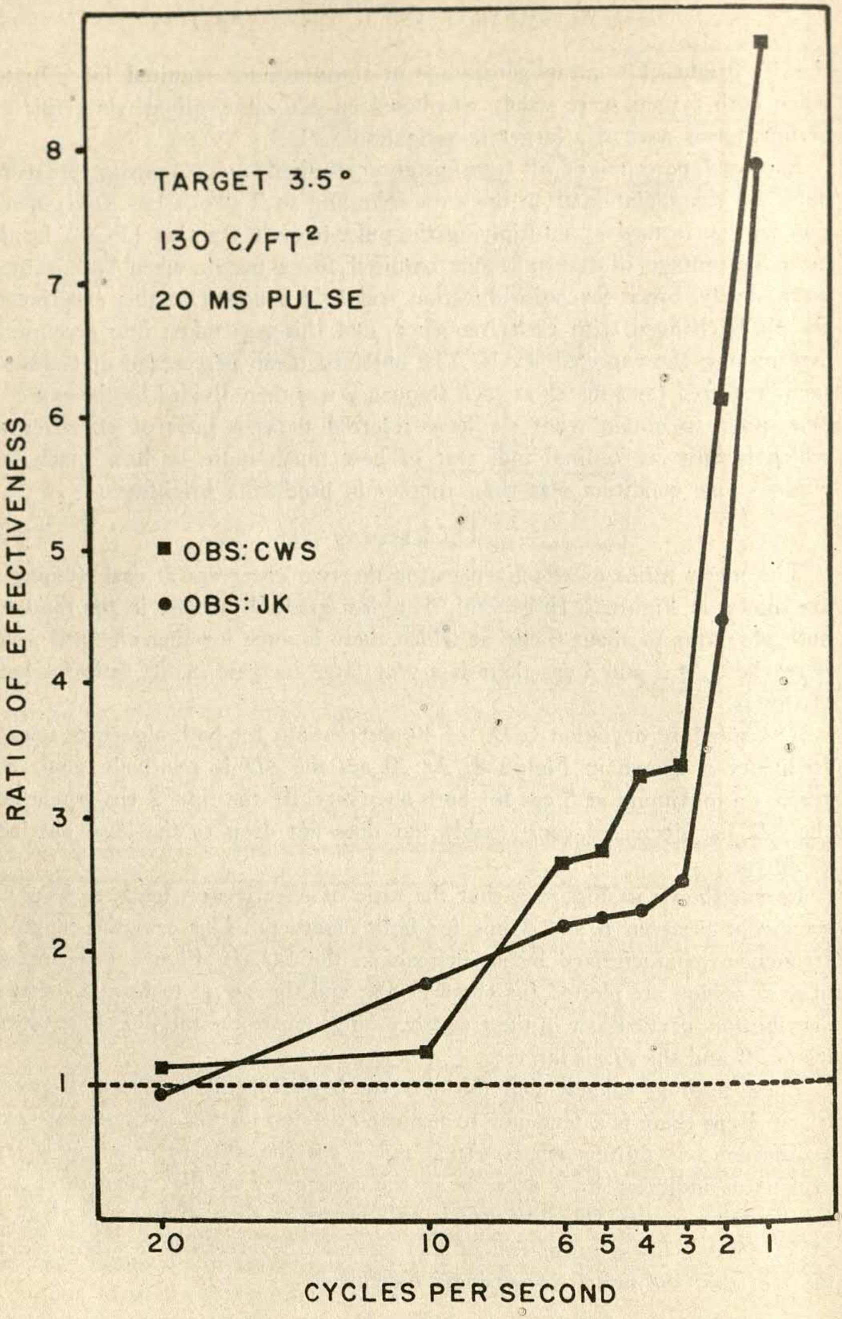


FIGURE 2
RATIOS OF EFFECTIVENESS OF A 20 MILLISECOND PULSE AT CERTAIN FREQUENCIES BETWEEN 20 AND 1 CPS, 130 C/FT^2 AND 3.5° TARGET, FOR TWO OBSERVERS

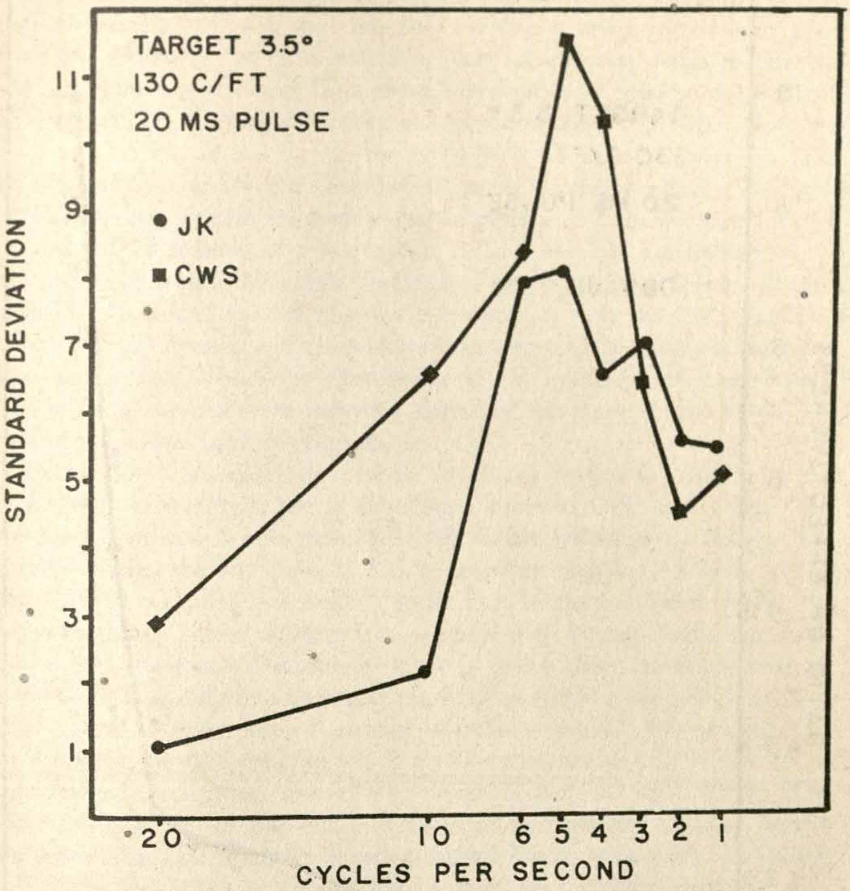


FIGURE 3
Standard Deviations for the Two Observers

D. Discussion

The findings of this investigation suggest that a shift in observer criterion occurs as the frequency of stimulation is decreased from 20 to 1 cps. Between 20 and 10 cps the observer is making his matches to an average between the light and dark phase, and the variances tend to be quite small. Below 10 cps, there is a transition zone which is reached somewhere around 6 cps and extends down to 3 cps. During this transitional period, the observer vascillates between matching to an average at one time and to the flash at another as

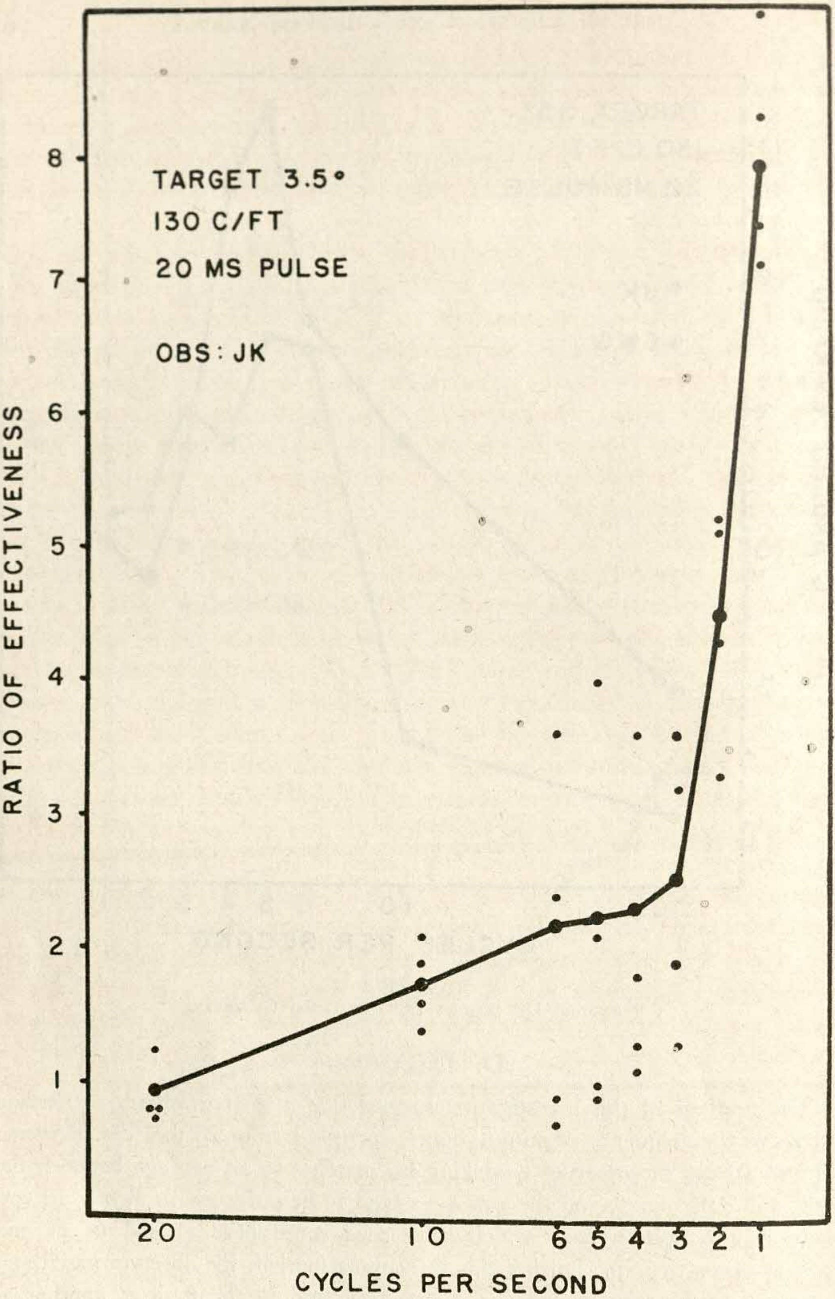


FIGURE 4
Distribution of Session Means (10 Trials) Around the Grand Mean for One of the Observers

indicated by the points plotted in the scattergram. It is interesting to note that the shift in criterion does not occur within a given session but rather from one session to the next, indicating that the observer tends to establish and maintain his criterion throughout any particular session but remains unaware of the distinction between the two operations (criteria), and this leads to a failure on the part of the investigator to make a distinction. The investigator also makes the same failure when he observes under the same conditions. The transitional zone is characterized by a large increase in variance and the formation of what appears to be a bimodal distribution in the percentage of transmission (the luminance) required for a match. At frequencies of 2 and 1 cps the observer can match only to the flash, and thus a new criterion level is reached which is characterized by a decrease in variance and a large increase in effectiveness of the intermittent input. Verbal reports from the observers indicated that they were not aware of the occurrence of a shift in their criterion.

Ranney (26) has carefully defined six classes in the appearance of the target with intermittent photic stimulation between CFF and 1 cps. The frequencies utilized in this experiment fall within the range of three of the discrete classes she has defined: i.e., a series of flashes (1-5 cps), coarse flicker (5-10 cps), and fine flicker (10-20 cps). In the fine flicker region, the target brightness is best characterized as a field with a large steady component and a fluctuation superimposed on it. It is under these conditions that an observer finds it difficult to separate the light and dark phase and, therefore, must match to some sort of average effect between the two phases. The transitional zone (3-8 cps), in which the observer may or may not match to an average, encompasses part of the coarse flicker and flash range. The observer can match to an average in this zone if instructed to do so, but in the unstructured situation his criterion shifts between averaging or matching only to the flash. If he matches only to the flash, the obtained brightness levels will be higher than those obtained with fine flicker; but if he is attempting to match to an average, the apparent brightness will be below that obtained with fine flicker as is indicated in Figure 3. This is not too surprising when we consider that the ratio of the dark to the light phase is much larger in this zone than it is with fine flicker. Under the two or single flash condition, the observer can only match to the flash, and the obtained brightness increases sharply. In light of what has been said above, we would expect the apparent brightness to be greatly reduced if the observer could match to an average at this point, but this sort of operation would be inappropriate even if it were possible. At the high frequencies, matching to an average between the light and the dark phase is compelling because the limitations of the visual mechanism make it impossible for the observer to distinguish clearly between the two phases. However, with the slow frequencies the situation is reversed.

If the observer is utilizing a particular criterion under one set of conditions and then finds it necessary to shift his criterion because of a change in the stimulus conditions, then it would seem logical to assume that the operations he is performing are related to a different phenomenon. Under such conditions, it would be an error to treat all of the data as if they were descriptive of a single phenomenon.

In light of these findings, it is suggested that *brightness enhancement* as Bartley (4) originally defined it be limited to describing the brightness phenomenon that occurs under those frequencies of intermittent stimulation that result in a brightness stable enough to be compared to a steady target. The brightness levels obtained with input rates producing a series of discrete flashes are more appropriately dealt with as single pulse phenomena which, as Boynton suggested, are not productive of enhancement phenomena under the original definition of enhancement.

Various additional experiments, in which aspects of the phenomenology not touched upon here are dealt with, can be designed to demonstrate this proposal more fully.

E. SUMMARY

The present paper points out the fact that as the rate of intermittent photic stimulation is reduced from CFF to the point at which each member of the train of pulses is greatly enough separated from the others temporally to function as an isolated single pulse, the target changes its appearance in a marked way. This whole span can be divided into several ranges. All frequencies above, let us say, approximately 8 or 10 pulses per second, produce fields which possess either a perfectly steady brightness or a steady brightness component predominant enough to enable comparison in brightness with a field produced by a continuous, uniform photic input. But, as frequencies drop below the range just mentioned, various changes occur in what is seen. These changes disenable the kind of direct comparison just indicated. From this point on, as frequency is reduced, the observer uses either the bright phase of the cycle or attempts to "average" the values of the two phases, in attempting to find some brightness in the intermittent field to match to the steady field. This "averaging" is not a direct sensory observation but a kind of judgment, and thus at low frequencies a vascillation between two operations occurs.

This paper also discusses the consequences of not recognizing the fact that more than one major phenomenon is involved in the intermittency span. Such consequences include the misinterpretation of brightness results and the unfounded criticism of the brightness enhancement theory put forth to account for brightness enhancement and related phenomena.

REFERENCES

1. ALPERN, M., & SUGIYAMA, S. Photic driving of the critical flicker frequency. *J. Opt. Soc. Amer.*, 1961, **51**, 1379-1385.
2. BALL, R. J. An investigation of chromatic brightness enhancement tendencies. *Amer. J. Optom. & Arch. Amer. Acad. Optom.*, 1964, **41**, 333-361.
3. BARTLEY, S. H. Temporal and spatial summation of extrinsic impulses with the intrinsic activity of the cortex. *J. Cell. & Comp. Physiol.*, 1936, **8**, 41-62.
4. ———. Subjective brightness in relation to flash rate and LDR. *J. Exper. Psychol.*, 1938, **23**, 313-319.
5. ———. The relation between cortical response to visual stimulation and change in the alpha rhythm. *J. Exper. Psychol.*, 1940, **27**, 624-639.
6. ———. Intermittent photic stimulation at marginal intensity levels. *J. of Psychol.*, 1951, **32**, 217-223.
7. ———. Some facts and concepts regarding the neurophysiology of the optic pathway. *Arch. Ophthal.*, 1958, **60**(2), 775-791.
8. ———. A clarification of some of the procedures and concepts involved in dealing with the optic pathway. In R. Jung and H. Kornhuber (Eds.), *The Visual System: Neurophysiology and Psychophysics.* Heidelburg: Springer-Verlag, 1961.
9. BARTLEY, S. H., & BISHOP, G. H. The cortical response to stimulation of the optic nerve in the rabbit. *Amer. J. Physiol.*, 1933, **103**, 159-172.
10. ———. Factors determining the form of the electrical response from the optic cortex of the rabbit. *Amer. J. Physiol.*, 1933, **103**, 173-184.
11. BARTLEY, S. H., & NELSON, T. M. Some relations between sensory end results and neural activity in the optic pathway. *J. of Psychol.*, 1963, **55**, 121-143.
12. BARTLEY, S. H., NELSON, T. M., & SOULES, E. M. Visual acuity under conditions of intermittent illumination productive of paradoxical brightness. *J. of Psychol.*, 1963, **55**, 153-163.
13. BARTLEY, S. H., PACZEWITZ, G., & VALSI, E. Brightness enhancement and the stimulus cycle. *J. of Psychol.*, 1957, **43**, 187-193.
14. BATTIG, W., VOSS, J., & BROGDEN, W. The effect of frequency of intermittence upon perceived brightness. *J. Comp. & Physiol. Psychol.*, 1956, **50**, 61-63.
15. BISHOP, G. H. Cyclic changes in the excitability of the optic pathway of the rabbit. *Amer. J. Psychol.*, 1933, **103**, 213-224.
16. BOURASSA, C. M., & BARTLEY, S. H. Some observations of the manipulation of visual acuity by varying the rate of intermittent stimulation. *J. of Psychol.*, 1965, **59**, 319-328.
17. BOYNTON, R. M. Some temporal factors in vision. In W. A. Rosenblith (Ed.), *Sensory Communication.* New York: M.I.T. Press, Wiley, 1959.
18. BROCA, A., & SULZER, D. La sensation lumineuse en fonction des temps. *J. Physiol. Path. Gen.*, 1902, **4**, 632-640.
19. ———. La sensation lumineuse en fonction des temps. *C. R. Acad. Sci.*, 1902, **134**, 831-834.

20. Colgan, C. M. The effect of observational technique on brightness enhancement. *Amer. J. Psychol.,* 1965, **78**, 471-475.

21. Gulick, W. L. Brightness enhancement with microsecond pulses. *Psychol. Rec.,* 1961, **11**, 249-256.

22. Halstead, W. C. A note on the Bartley effect in the estimation of equivalent brightness. *J. Exper. Psychol.,* 1941, **28**, 524-528.

23. Katz, M. S. Brief flash brightness. *Vision Res.,* 1964, **4**, 361-373.

24. Lindsley, D. M. Some facts and concepts regarding the neurophysiology of the optic pathway. (In discussion of article by S. H. Bartley.) *Arch. Ophthalmol.,* 1958, **60**(2), 775-791.

25. Nelson, T. M., Bartley, S. H., & Ford, Z. The brightness of part-spectrum targets stimulating the eye parafoveally at intermittency rates below fusion. *J. of Psychol.,* 1963, **55**, 387-396.

26. Ranney, J. E. Perceptions of subfusional photic intermittencies under various intensities and pulse-to-cycle fractions. Unpublished Doctoral dissertation, Michigan State University, East Lansing, Michigan, 1964.

27. Reese, T. W. The application of the theory of physical measurement to psychological magnitudes with three experimental examples. *Psychol. Monog.,* 1943, **55**, 1-89.

28. Reidmeister, C., & Grüsser, O. J. Flimmerlichtuntersuchungen an der Katzenretina, I. *Z. Biol.,* 1959, **111**, 241-253.

29. Schneider, C. W. An investigation of the influence of the interrelationship among certain variables and observer criterion on apparent brightness under conditions of intermittent stimulation. Unpublished Doctoral dissertation, Michigan State University, East Lansing, Michigan, 1965.

30. Taubman, R. E. Studies in judging number: I. Judgment of auditory number. *J. Gen. Psychol.,* 1950, **43**, 167-194.

31. ———. Studies in judging number: II. Judgment of visual number. *J. Gen. Psychol.,* 1950, **43**, 195-219.

Department of Psychology
Michigan State University
East Lansing, Michigan 48823

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 67-72.

SPECTROGRAPHIC ANALYSIS OF ELECTROENCEPHALOGRAMS UNDER CONDITIONS OF ALERTNESS AND RELAXATION*[1]

Department of Psychology, University of California, Los Angeles

J. A. GENGERELLI AND CHRISTOPHER E. PARKER

A. INTRODUCTION

The change in the human electroencephalogram (EEG) when the subject switches from a condition of serene relaxation to a condition of pronounced mental activity or active perception is well known (1, 7). Variously termed "alpha-blocking," "alerting reaction," or "desynchronization," it has been characterized in such terms as ". . . fast electrocortical activity which replaces the alpha rhythm when the blocking reaction is induced by opening the eyes . . ." (2, p. 893).

While such qualitative descriptions attempt to be faithful to what is seen on inspection of the EEG record under the two conditions, the tracings are too complex for unaided analysis (3, 4, 5, 6). Since EEGs may be justly compared to fairly well modulated noise, it is desirable to reduce them to Fourier transforms under the conditions of relaxation and alertness, respectively, in the same subject. An examination of the resulting spectrograms will then readily reveal any shifts in the distribution of energy among the component frequencies constituting the original signal.

B. METHOD AND APPARATUS

The subject reclined in a comfortable position in an electrically shielded room, with conventional needle electrodes inserted in the scalp in the occipital region and placed symmetrically on either side of the midline. After amplification by means of two Tektronics Type 122 preamplifiers arranged in series, the signal was fed simultaneously into a Model SS-5 Spectrum Analyzer and a Model 5D Grass Polygraph. The spectrum analyzer[2] used was essentially a narrow-band superheterodyne receiver which was repeatedly swept at

---

* Received in the Editorial Office, Provincetown, Massachusetts, on March 28, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This study was supported by a grant from the Faculty Research Committee of the University of California, Los Angeles.

[2] Probescope Company, Inc., Port Washington, N.Y.

predetermined intervals over a selected portion of the frequency band. At the same time, the horizontal deflection of the spot on a cathode ray tube moved in synchronism with the sweep. The vertical deflection of the spot was proportional to the output voltage of the analyzer. The resultant display was a plot of amplitude *versus* frequency over the frequency band of interest. In the experiment reported here the frequency band employed was between zero and 30 cycles per second. In preliminary experiments, a band between zero and 60 cps was utilized, but it was obvious that relatively little energy was to be found in Fourier components above 30 cps.

The spectrum analyzer displays the spectrogram on its cathode ray tube and simultaneously transmits it through a separate channel to any suitable recorder. We used two types of recorders. One was a CAT[3] coupled to an *X*-*Y* Plotter. This permitted averaging a number of consecutive sweeps and plotting out the result in a single graph. The other recorder, the Grass polygraph, was used for a separate recording of each successive spectrogram so that a record of the variability of the spectrograms from sweep to sweep could be secured. At the same time the EEG signal being analyzed was recorded on a separate channel (Figure 1).

For the frequency band between zero and 30 cps, it was necessary to set the sweep period at 8.4 seconds. Thus, in our experiment, each frequency component could not be measured at intervals shorter than 8.4 seconds. With this setting the resolving power of the analyzer was approximately three cycles.[4] As was expected, the spectrogram for a given sweep was not identical to that of the next sweep, although the general features were preserved (Figure 1). The plot-out from the CAT represents an average of 20 sweeps (Figure 2).

Fourteen subjects, males and females, were tested. After a period of several minutes during which the subject adjusted himself to the chair, he was told to close his eyes and to relax as completely as possible; he was, however, to

---

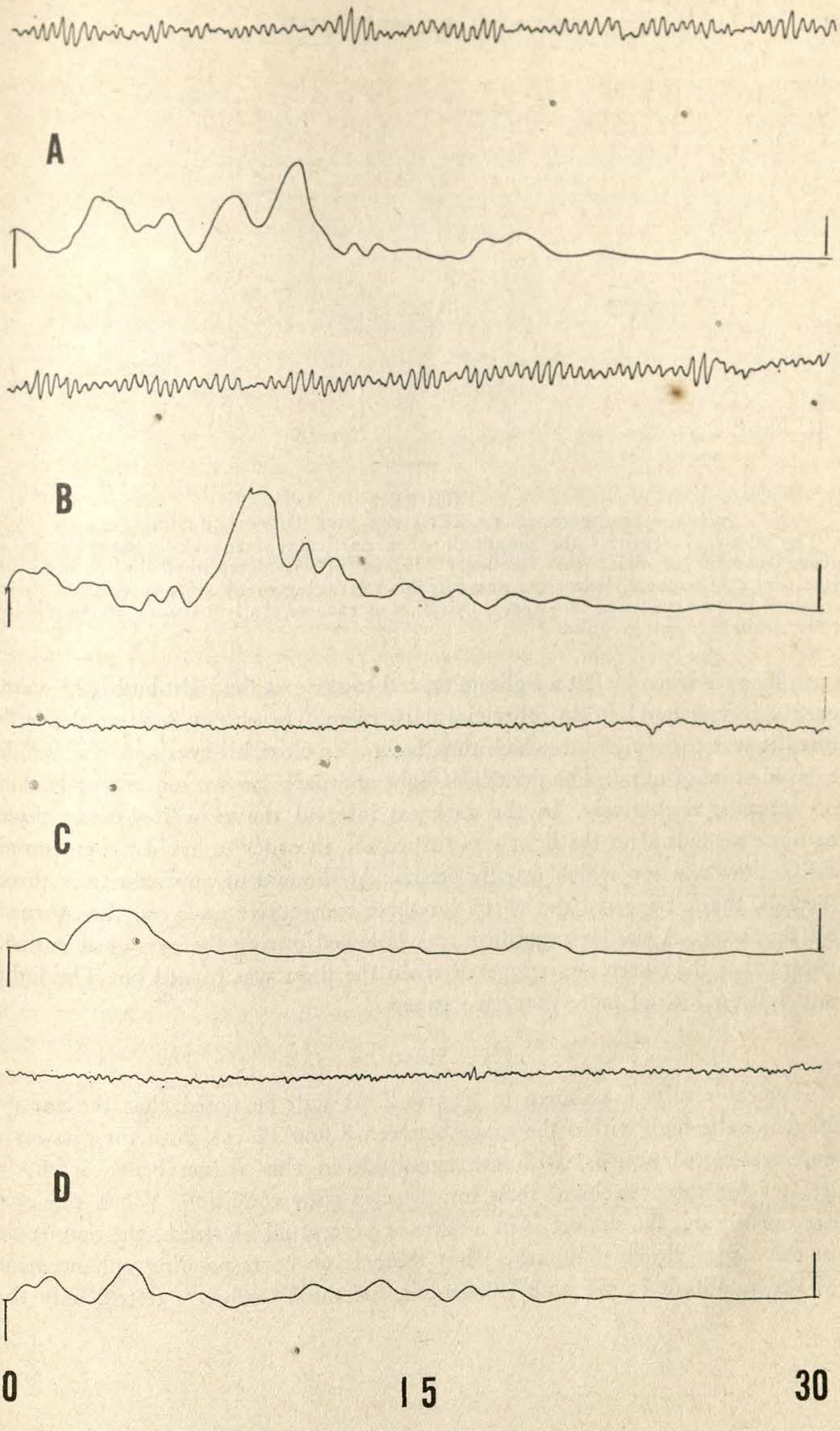
[3] Computer of Average Transients. Mnemotron Corporation, Pearl River, N.Y.

[4] The three parameters—frequency-band, sweep period, and resolving power—are not independent in this analyzer.

FIGURE 1

SIMULTANEOUS RECORDINGS OF THE EEG AND ITS SPECTROGRAM

A and B are successive sweeps of the analyzer with the eyes closed under condition of relaxation. C and D are successive sweeps with the eyes open and fixating the light. Note differences between A and B, although B follows immediately upon A. Similarly for C and D. The four recordings are from the same subject. Duration of each sweep is 8.4 seconds. The mark below the spectrographic tracing indicates the beginning; the one above indicates the end of a sweep.


A
B
C
D
0
15
30
CPS

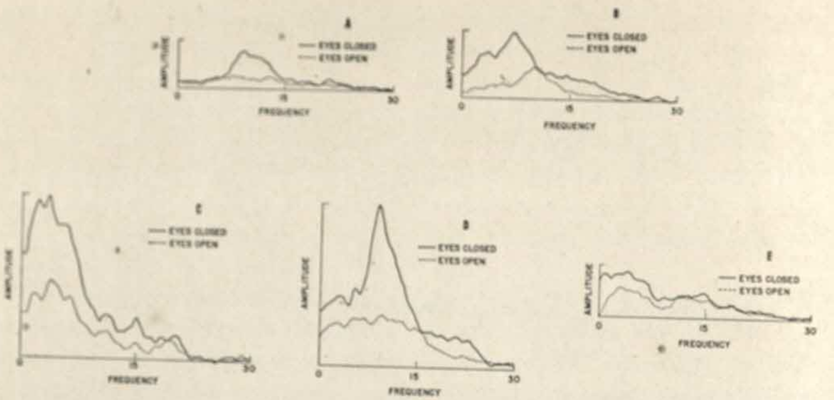


FIGURE 2

Average Spectrograms of EEGs for Five Different Subjects

The analyzer registers the amplitude of a particular frequency component at a given moment of time; the frequency "looked at" is determined by position of sweep at that moment. Hence the necessity for averaging many sweeps. Each of these tracings is the average of 20 sweeps. Duration of each sweep is 8.4 seconds. Amplitude is in arbitrary voltage units.

open his eyes when he felt a light go on and to stare at the light bulb (75 watt, ca. six feet distant) in an otherwise dark room. He was to look at the bulb until it was turned off, at which time he was to close his eyes and resume his relaxed state of mind. The periods of light and darkness were of about 18 and 35 seconds, respectively. In the darkness interval the *E* waited some three or more seconds after the light was turned off, in order to avoid the enhanced EEG slow reaction which usually occurs. At the end of approximately three seconds, the *E* triggered the sweep for three consecutive analyses, then turned off the sweep. A similar procedure was followed during the eyes-open period, except that the sweep was triggered when the light was turned on. The light was left on for not more than two sweeps.

C. Results

Typical results are shown in Figure 2. It will be noted that the energy level is quite high within the range between 8 and 12 cps, both for eyes-open and eyes-closed periods, with the amplitude in this region being definitely greater for the eyes-closed than for the eyes-open condition. When the eyes are opened and the subject is in a state of perceptual alertness, the amplitude in this range clearly diminishes; but there is no corresponding enhancement of the amplitude in the high frequencies. Of the 14 subjects tested, only the

S whose spectrogram is shown in Figure 2B gave a shift in amplitude toward a slightly increased frequency. The other subjects either showed a small diminution in amplitude above 15 cps or no change above this point.

It will be noted that considerable individual differences are to be found among the average spectrograms of various subjects. In spite of these differences, however, the trend of the discrepancies between the eyes-closed and the eyes-open conditions was maintained in each of the 14 subjects, although the size of the discrepancies showed considerable variation. Forty sweeps were averaged for two subjects with similar results, the only notable difference being decreased "undulations" in the spectrograms.

D. Discussion

When EEGs obtained in a relaxed state are compared visually with those registered during periods of mental or perceptual alertness, one gains the impression that the high frequency, low amplitude waves are more prominent in the latter than in the former. The Fourier transforms found in the present experiment show, however, that there is no energy gain in frequency components anywhere in the range between zero and 30 cps. There is a marked diminution in amplitude between zero and approximately 15 cycles, and thereafter a very small decrease, if any. We are therefore forced to conclude that high frequency components exist in EEG tracings obtained in the relaxed state, but are not apparent because of the greatly increased amplitude of the low frequency components.

E. Summary

Analyses of human EEGs were made with a spectrum analyzer under conditions of perceptual alertness and of relaxation. The range of frequency components studied was from zero to 30 cps. Comparison of the spectrograms obtained under the two conditions showed a marked diminution in the amplitude of frequency components below 15 cps, but no increases anywhere in the spectrum. In some subjects there was diminution in amplitude in the high frequency components also, but this change was quite small.

References

1. Berger, H. Uber das elecktrenkephalogram des menschen. *Arch. Psychiat. Nervenkr.*, 1929, **87**, 527-570.
2. Calvet, J., Calvet, M. C., & Langlois, J. M. Diffuse cortical activation waves during so-called desynchronized EEG patterns. *J. Neurophysiol.*, 1965, **28**, 893-907.

3. Eidelberg, E., & Cheshire, F. C. Spectrum analyzer for EEG. *Electroenceph. & Clin. Neurophysiol.*, 1965, **18**, 85-87.
4. Grass, A. M., & Gibbs, F. A. A Fourier transform of the electroencephalogram. *J. Neurophysiol.*, 1938, **1**, 521-526.
5. Hord, D. J., Johnson, L. C., Lubin, H., & Austin, M. T. Resolution and stability in the autospectra of EEG. *Electroenceph. & Clin. Neurophysiol.*, 1965, **19**, 305-308.
6. Knott, J. R., Gibbs, F. A., & Henry, C. E. Fourier transforms of the electroencephalogram during sleep. *J. Exper. Psychol.*, 1942, **31**, 465-477.
7. Morgan, C. T. Physiological Psychology (3rd ed.). New York: McGraw-Hill, 1965.

Department of Psychology
University of California, Los Angeles
Los Angeles, California 90024

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 73-81.

THE INDUCTION OF ANCHORING EFFECTS IN ABSOLUTE JUDGEMENTS THROUGH DIFFERENTIAL REINFORCEMENT*

Department of Child Development and Family Relationships, Cornell University

WILLIAM G. LE FURGY

A. INTRODUCTION

Numerous studies have shown that the introduction of an anchor stimulus that lies a reasonable distance beyond the range of stimuli to be judged results in the expansion of the categories of the *S*'s response scale in the direction of the anchor [see Cohen (1), Hunt and Volkman (2), McGarvey (6), Postman and Miller (7), and Rogers (8)]. Analogous studies have shown that such factors as the *S*'s prior experience [see Larimer and White (4), and Tresselt (11)], personal characteristics [see Marks (5)], and social attitudes [see Sherif, Taub and Hovland (10)] can serve as anchors that produce systematic influences in the scaling of relevant stimuli.

The present experiment is designed to induce such anchor effects by experimentally manipulating the quality of the *S*'s affective experience with stimuli that are then scaled by the absolute method. Traditionally, the procedures of classical psychophysics have sought to eliminate or to train out individual differences by presenting a few *S*s with a small number of stimuli (usually no more than seven), and by seeking discriminations of relatively small physical differences over a large number of trials. Other investigators have attempted to maximize the contribution of subjective factors in judgement by presenting stimuli to the *S* under ambiguous conditions. Neither of these procedures necessarily represents the situation as it occurs in the everyday experience of the *S*. Outside of the psychological laboratory relatively few of the stimuli that impinge upon our sensory systems do so at extremely low intensities or for durations measured in milliseconds. To be sure, the differences between these stimuli are more often than not extremely small, but at any given moment they number not in the tens but in the thousands. It would seem logical to assume that a stimulus situation that presents the *S* with a surplus of potential information can be regarded as "unstructured"

---

* Received in the Editorial Office, Provincetown, Massachusetts, on March 28, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

as a situation in which the stimuli are impoverished. When the *S* is confronted with a stimulus situation that is physically ambiguous, the effect is to increase the number of possible responses to one or to a few stimuli. In a situation where a surplus of information is provided, the net effect is to increase the number of stimuli that are to be assigned to a limited number of response designations. Both of these situations offer the *S* a relatively large number of response alternatives. It might be expected that in choosing among such a variety of alternatives the *S* will utilize those aspects of his past experience that are, for him, most relevant.

B. METHOD

1. *Subjects*

The *S*s were 30 female undergraduates who volunteered for the experiment. The *S*s were assigned to the experimental and control groups on a random basis.

2. *Apparatus*

The apparatus in which the stimuli were displayed was a panel of Masonite board approximately two by four feet in dimension. The face of the panel was composed of three vertically sliding doors which, through a string and pulley arrangement, could be operated by either the *S* or the *E*. Three identical sets of 15 stimulus cards were mounted in a rack behind the panel. The 45 stimulus cards were arranged in the rack so as to permit the *E* to vary the presentation order of the 15 *different* stimuli, while at the same time allowing him to present *identical* stimuli behind each of the three doors on any given trial.

The stimuli consisted of a continuum of 15 circles ranging from 75.79 to 131.80 mm in diameter. The stimuli were drawn in India ink and centered on 11-inch squares of white poster board.

3. *Procedure*

The procedure for all *S*s consisted of a training session and a subsequent testing session. During the training session a series of 15 circles of increasingly different diameters were exposed to the *S*, five times each for a total of 75 trials. On each trial the *E* informed the *S* as to the numerical designation of the stimulus that was displayed. Subsequently, during the testing session these same 15 original stimuli plus 14 intermediate stimuli (i.e., circles whose diameters were intermediate to the original stimuli) were presented to the *S*. The *S* was then instructed to judge them in accord with the designations learned in the training session. The distinction between the two experimental groups and the control group is to be found in the conditions of their respective training sessions. The instructions given to the *experimental Ss* prior to the training session were as follows:

> This is an experiment to see how quickly people are able to perceive relationships. In order to make it interesting we are going to play kind of a game. You will keep score here with this pad and pencil. Behind these doors are 15 circles of various sizes. Each circle is worth a different number of points. There are three different kinds of circles: those that add to your score by giving you positive points; those that subtract from your score by giving you negative points; and some that are neutral, that do not affect your score in any way. You will discover that the score values of the circles are somehow related to their size. Your job will be to try to figure out the sequence or sequences of order and position being used so that you can open the doors in such a way so as to amass the highest number of positive points. Once you "catch on" to the pattern this should not be difficult.

The *S* was then seated five feet away from the display panel and shown how to operate the doors. The *S* was shown how to keep a cumulative algebraic score by means of which she was aware of her standing throughout the procedure.

For Experimental Condition 1, the 15 stimuli were given the following numerical designations. The five smallest stimuli (75.79 to 88.80 mm in diameter) were assigned values of — 100, — 80, — 60, — 40, and — 20 respectively. The five middle-sized circles (92.39 to 108.15 mm in diameter) were assigned values of N (neutral) 100, N80, N60, N40, and N20 respectively. The largest stimuli (112.60 to 131.80 mm in diameter) were assigned values of + 20, + 40, + 60, + 80, and + 100 respectively. Thus, under Experimental Condition 1 the *smallest* stimulus had the *highest negative* valence and the *largest* stimulus had the highest *positive* valence with the remainder distributed between.

For Experimental Condition 2, all procedures were identical to those in Condition 1 except for the fact that the signs of the numerical designations were reversed so as to give the *smallest* circle the highest *positive* valence and the *largest* circle the highest *negative* valence.

During the experimental procedures, on each trial the *S* was required to select and open a door and thus present herself with the given stimulus. After the *S* had raised a door, the *E* immediately announced the score value of the stimulus to the *S*. The *S* then applied this value to her cumulative total.

From her point of view, based on the information that she had been given, the *S* believed that on any given trial she was making a choice, possibly, between three different-sized stimuli. In actuality, the order and frequency with which the *S*s encountered the various stimuli had been completely predetermined by the *E*. This was accomplished by placing identical stimuli behind all three doors on each trial. In this manner all *S*s were exposed to each of the 15 stimuli five times in a predetermined, though random, order.

For the Control Group the order and the frequency of the presentation of the stimuli were the same as under Experimental Conditions 1 and 2, but several modifications in procedure and instructions were introduced so as to remove any affective connotation from the situation.

Under this condition the role of the *S* was a relatively passive one. The instructions for the control group were as follows:

> Behind these doors is a series of circles of various sizes. To these circles I have assigned a numbering system. (The numbering system under this condition was identical with that of Experimental Conditions 1 and 2 except for the fact that the plus and minus symbols were replaced by the letters A and B respectively.) I am going to open the doors one at a time. Every time I show you a circle I will tell you its number. There are three different kinds of circles, A, B, and N. As we proceed you will see that the numbering system is somehow related to the size of the circle. All I want you to do is to attend to the circle and the number that I give it.

The *E* opened the doors of the display panel in random order. On the third trial the *S* was informed that the presentation order was completely random; otherwise the presentation was identical with that used under Experimental Conditions 1 and 2. On the average, training sessions for both control and experimental *S*s lasted from 30 to 35 minutes.

The testing sessions were identical for both experimental and control *S*s. Immediately after completion of the training session, the *S*s proceeded to the testing apparatus.

The testing apparatus, which was in a different room from the experimental apparatus, consisted of a table on which was mounted a screen with an opening similar in dimensions to the doorways of the display panel used in the training session. The *S*'s position relative to the screen was the same as in the training session. Upon being tested the *S* was given the following instructions.

> Now I am going to show you the same circles and ask *you* to tell *me* the score values.

Unknown to the *S* she was to be presented with not only the original 15 stimuli, but also 14 stimuli of "intermediate" sizes. Between each of the original circles on the scale there was now inserted an additional circle that was intermediate in size to the two adjacent original circles. As a result of the insertion of the intermediate stimuli, the *S* was now to judge a scale of 29 rather than 15 different stimuli. This presentation differed from the presentation used in the training session in that the random order was different and the stimuli were presented in blocks of 29 trials each. Each block of 29 presentations included all of the 29 stimuli in a different random order. A total of three blocks of 29 trials were presented. Care was taken not to make it explicit to the *S* that each of the blocks contained the entire scale of 29 stimuli. All judgments were made as in the method of single stimuli and every effort was made to eliminate any feedback or knowledge of results to the *S*.

At the conclusion of the testing session the *S*s were questioned as to any introspections that they might have. The experimental *S*s were specifically questioned concerning any affective arousal that they might have experienced in association with the positive and the negative stimuli.

C. RESULTS

Figure 1 gives the median judgement of stimulus size for each of the three segments of the stimulus scale under each condition of the experiment. This figure shows the principal result of the experiment to be the relative distortion of judgements in favor of the positively reinforced segments of the scale. Despite the fact that a negative time order effect produced a general overestimation of stimulus size under all conditions of the experiment, Experimental Group 1, for whom the largest stimuli were positive and the smaller stimuli negative, shows consistently greater overestimations of stimulus size than does Experimental Group 2, for whom these affective connotations were reversed. An analysis of variance of the sum of overestimation minus the sum of underestimation of stimulus size for each *S* by experimental condition yields a significant main effect of experimental condition at the .005 level. A multiple comparison of the three means, carried out by a procedure described by Ryan (9), shows that the difference between the Control Group and Experimental Group 2 is not significant. All other comparisons are significant with $p < .05$.

The shift in scale values toward the positively reinforced segment of the continuum is most clearly shown in Figure 2. Figure 2 presents category limens for all *S*s for each experimental condition. These limens were computed by a method described by Johnson (3, pp. 330-337), and are essentially equivalent to the more familiar concept of Points of Subjective Equality. Each

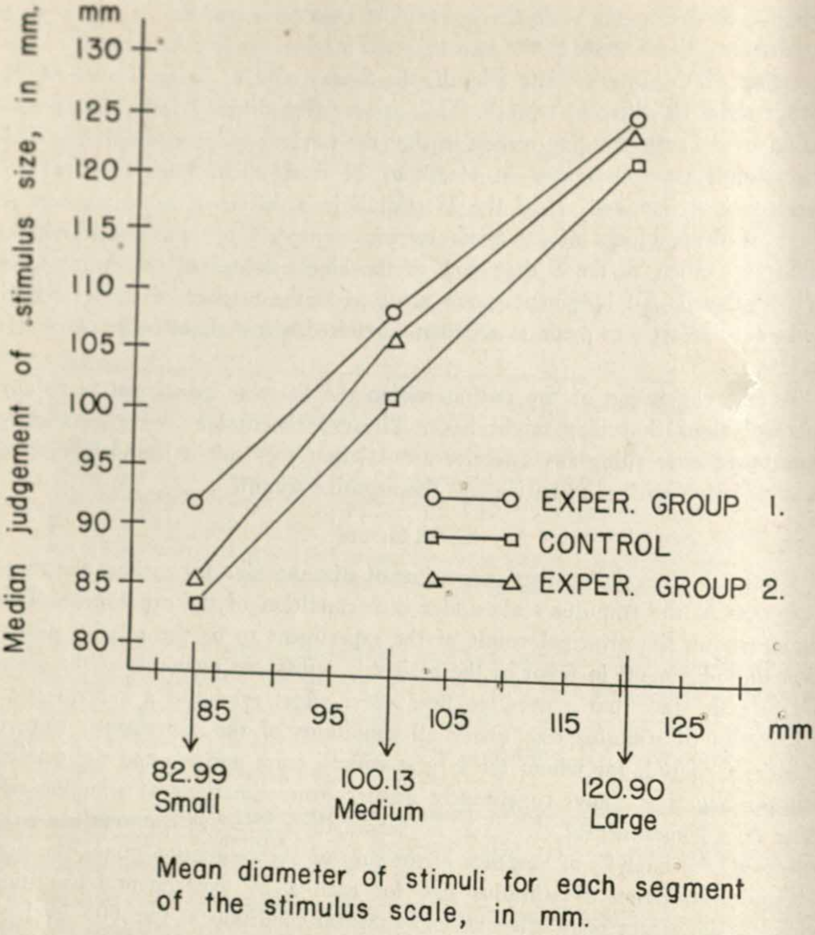


FIGURE 1

Median Judgement of Stimulus Size for Each Condition of the Experiment

limen represents a Least Squares solution to the function derived by plotting the cumulated proportion of judgements "greater than" each category, by response category. Figure 2 shows clearly that it is the *entire* response scale that shifts as the result of differential reinforcement and, further, that this shift is in the direction of the positively reinforced segment of the continuum.

This is reflected in Figure 2 by the general expansion of the category thresholds under the two experimental conditions in the direction of the negatively reinforced segments of the scale. In both experimental groups that portion of the continuum responded to as positive constitutes more than 41 per cent of the total stimulus scale while that portion of the continuum responded to as negative is less than 28 per cent of the total scale. In each case the *S*s tended to make their judgements "as if" an anchor had been introduced beyond the end of the negatively reinforced segment of the scale.

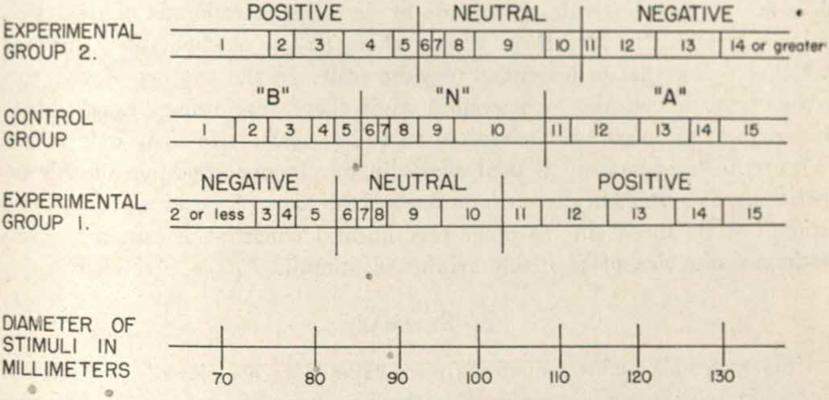


FIGURE 2

Category Limens for All Subjects for Each Experimental Condition

At the conclusion of the experiment, all of the *S*s responded in the affirmative when they were asked: "When we were in the other room did you feel better about finding big circles than small circles (or *vice versa* depending upon the experimental condition)?" In most instances this question was superfluous, for the *S*s' involvement with the training task was quite apparent. This was undoubtedly facilitated by the fact that the *S*s were all students of the *E*, and were performing in the *E*'s presence on an impossible task that ". . . should not be too difficult once you catch on."

The difficulty of the task presented to the *S*s during the testing session was attested to by the *S*s' comments to this effect. When large circles had been reinforced, it was not unusual for an *S* to inquire of the *E* why he had not shown her any small circles, or *vice versa* if small circles had been positively reinforced. The *S*s resorted to a variety of devices in attempts to bring some stability to their response scales. The most interesting of these occurred when several *Ss* attempted to anchor previous judgements by joining their thumb and forefinger in the rough outline of circle as if literally to "hold on" to their scale.

D. Discussion

These results seem to leave little doubt that the specific quality of prior affective experience has a significant impact upon how subsequent judgements of these stimuli are organized. Specifically, the association of affect with stimuli does not necessarily result in increases or decreases in estimates of their size. Rather it serves to give them "figure character" as anchor stimuli around which the *S* organizes his entire response scale. In the context of this particular experiment, the *Ss* responded as if they were using "exaggerated" instances of the negatively reinforced stimuli as anchors in their judgements. The term "exaggerated" is used advisedly, for in one condition of this experiment *Ss* systematically underestimated the size of positively reinforced stimuli while the *Ss* in the other experimental condition consistently overestimated the size of positively reinforced stimuli.

E. Summary

Thirty female undergraduate *Ss* were exposed to a series of 15 circles of increasingly different diameters. For the two experimental groups, the five circles at each end of the continuum were given affective connotations through a process of differential reinforcement. Subsequently, the 15 original stimuli plus 14 intermediate stimuli were judged by the method of single stimuli. The results showed that, relative to a control group, the experimental *Ss* tended to shift their response scales in the direction of the positively reinforced segment of the continuum, with regard to both response frequency and size estimation.

References

1. Cohen, N. E. The relativity of absolute judgments. *Amer. J. Psychol.*, 1937, **49**, 93-100.
2. Hunt, W. A., & Volkman, J. The anchoring of an affective scale. *Amer. J. Psychol.*, 1937, **49**, 88-92.
3. Johnson, D. M. The psychology of thought and judgment. New York: Harper, 1955.
4. Larimer, G. S., & White, B. J. Some effects of momentary reward and knowledge of results on judgment. *J. Exper. Psychol.*, 1964, **67**, 27-32.
5. Marks, E. Skin color judgments of Negro college students. *J. Abn. & Soc. Psychol.*, 1940, **22**, 3-14.

6. McGarvey, H. R. Anchoring effects in the absolute judgment of verbal materials. *Arch. Psychol.*, 1943, **39**, No. 281.

7. Postman, L., & Miller, G. A. Anchoring of temporal judgments. *Amer. J. Psychol.*, 1945, **58**, 43-53.

8. Rogers, S. The anchoring of absolute judgments. *Arch. Psychol. (New York)*, 1941, **37**, No. 261.

9. Ryan, T. A. Multiple comparisons in psychological research. *Psychol. Bull.*, 1959, **56**, 26-47.

10. Sherif, M., Taub, D., & Hovland, C. I. Assimilation and contrast effects of anchoring effects of anchoring stimuli on judgments. *J. Exper. Psychol.*, 1958, **55**, 150-155.

11. Tresselt, M. E. The effect of the experience of contrasted groups upon the formation of a new scale of judgment. *J. Soc. Psychol.*, 1948, **27**, 209-216.

Department of Child Development and Family Relationships
New York State College of Home Economics
Cornell University
Ithaca, New York 14850

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 83-88.

DEVELOPMENT AND APPLICATION OF AN OBSERVATIONAL METHOD: CONTINUING RELIABILITY TESTING*[1]

Regional Primate Research Center and Department of Psychiatry, University of Washington

RUTH A. BOBBITT, BETTY N. GORDON, AND GORDON D. JENSEN[2]

A. INTRODUCTION

Continuous demonstration of reliability, or agreement among observers, is a basic principle of the authors' previously described method for observing the behavior of mother and infant pigtail monkeys (2). It cannot be assumed that interobserver agreements, determined after intensive practice and before collection of data, will be maintained over a prolonged period. The authors have therefore suggested (1) that once intraobserver agreement is achieved tests of interobserver agreement should be scheduled throughout an investigation. This paper presents the results of such a study of interobserver agreement over two and one-half years, during which time data were collected on three groups of subjects observed in three different environments. Each group consisted of four mother-infant pairs of monkeys. They were observed during five random 10-minute samples (a 50-minute block) of behavior each week for 26 weeks. A pair of observers simultaneously scored the behavior of mother and infant in one of the five observations in each block.[3]

B. METHOD OF STUDY

The conservative reliability measure of Agreement Percentage (A.P.) (4) requires few assumptions with respect to validity and is a clear statement of the agreement among observers on the behavior that occurred.[4] In general, the

---

* Accepted by John E. Horrocks of the Editorial Board; received in the Editorial Office, Provincetown, Massachusetts, on March 28, 1966; and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This research was supported by a grant (HD 00883-03) from the National Institute of Child Health and Human Development, and by Public Health Service Grant No. FR00166.

[2] The authors wish to acknowledge the assistance of Kathryn S. Strother and Jaclyn H. Wolfheim in collecting and analyzing these data.

[3] Observers were required to attain a .75 Agreement Percentage criterion before collecting data.

[4] No. of agreements ÷ (No. of agreements + No. of disagreements + No. of omissions).

A.P. was derived for each of the following: (*a*) the observers as a group, (*b*) each observer in relation to the others, (*c*) each type of behavioral unit recorded, (*d*) the group of behaviors included in a given behavioral dimension, (*e*) the behavior of each member of a subject-pair, and (*f*) the overall behavior of all subjects observed in each group.

In the observational method discrete units of behavior, operationally defined, were recorded in their order of occurrence. The observer related each unit of behavior (there were approximately 45) to three referents so that, at each change in behavior, he recorded into a tape recorder the position of the animals relative to each other, the new unit of behavior, who did it, and to whom it was directed. The verbal system of coding was designed to have cue value to observers. A behavior was recorded when it began. For some behaviors, the termination was also recorded, but a behavior continuing in successive time intervals was not recorded for each interval. Thus, behaviors of long duration, that were easily identified, did not receive repeated agreement counts. Momentary behaviors that actually began more often than prolonged behaviors required immediate identification and had to be scored each time they occurred. Finally, since the behavior that did occur was the focus of study, agreements were not scored for the time intervals in which observers agreed that no recordable behavior had occurred.

In reliability testing, each behavioral event and each termination observed were scored as (*a*) an agreement, (*b*) a disagreement, or (*c*) an omission on the part of one of the observers. An A.P. was calculated for each observation. This score, together with an analysis of nonagreement, was made available to the observers for discussion and practice prior to the next reliability check. Discussions of disagreements were arbitrated and definitions of behavioral units were clarified by the senior author, who designed the method and was not involved in data collection. The A.P.'s were tabulated over the 26 weeks each pair of animals was studied. Summaries were obtained for each observer with other observers, each animal pair, animal groups, and all animals. A.P. was also summarized by dimension.

C. RESULTS

Table 1 summarizes the A.P.'s and the actual number of agreements, omissions, and disagreements recorded for all behavioral dimensions and groups of subjects. It shows that observers were able to maintain approximately the same level of agreement for all groups of subjects, but that their A.P.'s differed somewhat over dimensions. Lowest agreement was on behavioral events in the visual dimension. It should be noted, however, that in this

TABLE 1

NUMBER OF AGREEMENTS, OMISSIONS, AND DISAGREEMENTS, AND THE AGREEMENT PERCENTAGE (A.P.) FOR ALL DIMENSIONS AND GROUPS

Variables	Position	Posture	Locomotion	Visual	Oral	Manipulation	Total
			Group I (N = 4)				
Total behavior	1984	4431	2456	1279	1666	2362	14,178
Agreements	1757	3730	1795	707	1337	1720	11,046
Omissions	218	606	358	475	301	562	2520
Disagreements	9	95	303	97	28	80	612
A.P.	.89	.84	.73	.55	.80	.73	.78
			Group II (N = 5)				
Total behavior	2014	6081	4178	1763	1920	3116	19,072
Agreements	1779	5064	3351	954	1558	2363	15,069
Omissions	227	858	517	664	319	634	3219
Disagreements	8	159	310	145	43	119	784
A.P.	.88	.83	.80	.54	.81	.76	.79
			Group III (N = 4)				
Total behavior	1234	2603	2212	998	1146	1748	9941
Agreements	1097	2071	1788	549	965	1321	7791
Omissions	126	444	255	388	158	362	1733
Disagreements	11	88	169	61	23	65	417
A.P.	.89	.80	.81	.55	.84	.76	.78
			All groups				
Total behavior	5232	13,115	8846	4040	4732	7226	43,191
Agreements	4633	10,865	6934	2210	3860	5404	33,906
Omissions	571	1908	1130	1527	778	1558	7472
Disagreements	28	342	782	303	94	264	1813
A.P.	.89	.83	.78	.55	.82	.75	.79

study only 4040 out of 43,191 (.094) of the behaviors recorded were in this dimension. Agreement in all other dimensions was above the acceptable .75 level.

From Table 2 it may be seen that essentially the same reliability level was maintained for all three groups of subjects by each pair of observers. Table 3 shows that the level of agreement between observer pairs did not vary essentially as the infants grew older and the number and the complexity of behaviors increased. This consistency is quite significant since the number of behaviors observed in a 10-minute period ranged from as low as 50 in the first week of life to over 400 later in life. Thus, essentially the same A.P. was maintained by all observer pairs over all animal groups throughout the study, and this level was above .75 through all but the relatively infrequent visual dimension.

TABLE 2
AGREEMENT PERCENTAGE FOR ALL OBSERVER PAIRS OVER THE THREE GROUPS OF SUBJECTS

Observer pairs	Groups I	II	III	Total
S & G	.82	.81	.80	.81
S & W	.79	.81	.79	.80
G & W	.79	.78	.79	.79

TABLE 3
AVERAGE AGREEMENT PERCENTAGE (A.P.) BY INFANTS' AGE (IN WEEKS)

Age in weeks	A.P.	Age in weeks	A.P.	Age in weeks	A.P.
1	.79	10	.78	19	.77
2	.78	11	.80	20	.82
3	.79	12	.78	21	.79
4	.78	13	.79	22	.80
5	.81	14	.81	23	.79
6	.80	15	.80	24	.79
7	.80	16	.77	25	.80
8	.81	17	.81	26	.82
9	.77	18	.79	Total	.79

Failure of observers to agree was a result of either disagreement over the proper name for the behavior or omission by one observer of a behavior which was scored by the other observer. Table 4 shows that omissions constituted a much higher percentage of nonagreements than did disagreements. Only in the locomotion dimension was disagreement responsible for a high proportion of the nonagreements. Thus, it can be said that most of the nonagreements were over whether or not a behavior did, in fact, occur. If both observers saw a behavior, they were quite unlikely to disagree on what it was. The lower part of Table 4 shows that each observer tended to be responsible for 40 to 60 per cent of the omissions scored for each observer pair, but that the percentage of omissions was generally highest for observer W, who was trained to criterion midway in the study. The fact that she tended to underscore in relation to the experienced observers suggests that observers continue to gain speed and completeness after achieving criterion.

TABLE 4
THE PERCENTAGE OF NONAGREEMENT IN EACH DIMENSION THAT IS OMISSION AND DISAGREEMENT, AND THE PERCENTAGE OF TOTAL NONAGREEMENT ATTRIBUTED TO EACH OBSERVER

	Dimensions					
	Position*	Posture	Locomotion	Visual	Oral	Manipulation
A.P.**	.89	.83	.78	.55	.82	.75
% Omissions	.95	.85	.59	.83	.89	.86
% Disagreements	.05	.15	.41	.17	.11	.14
Observers***						
$\frac{\text{Total S}}{(S+W)+(S+G)}$	.46	.42	.50	.50	.47	.52
$\frac{\text{Total W}}{(W+S)+(W+G)}$	.66	.65	.60	.47	.64	.64
$\frac{\text{Total G}}{(G+W)+(G+S)}$	.41	.51	.42	.52	.34	.38

* The position of the animals relative to each other: i.e., very close, close, near, separated.

** Agreement Percentage.

*** S, W and G = individual observers.

D. DISCUSSION

It is our belief that the reliability of behavioral data should be demonstrated throughout an investigation, and that such continuous reliability testing is an aid to the maintenance of agreement among observers. The results presented here show that, with this approach, agreement among several observers can be maintained over a prolonged period of behavioral development, with different subjects, and on most behavioral dimensions.

Finally, the level of agreement maintained during this study substantiates the claim (2) that a useful and objective method for securing detailed observations of ongoing behavior has been developed. This technique of monitoring interobserver agreement, with a modification of the observational method, has therefore been applied to larger groups of monkeys in our laboratory and to the study of human mothers and their preschool children (3).

E. SUMMARY

The agreement between pairs of observers using a verbal code to record the behavior of mother pigtailed monkeys and their infants during a developmental period was repeatedly tested. Agreement percentages were derived for each test and summarized across subjects, dimensions, observers, and time. Before beginning to record experimental data, each observer attained an agreement criterion of .75. The study shows that all observers maintained or exceeded their performance over a prolonged developmental period for all subjects on all behavioral dimensions but the visual, which was the dimension least frequently recorded. Omission by one observer of a behavioral unit noted by the other was the primary source of nonagreement, although some disagreements on the proper code for a behavior also occurred. There are indications that such continuous testing with discussion of the specific results of each reliability check sustains observer-attention on operational definitions of behavior and on comprehensive reporting.

REFERENCES

1. BOBBITT, R. A., JENSEN, G. D., & GORDON, B. N. Behavioral elements (taxonomy) for observing mother-infant-peer interaction in *Macaca nemestrina*. *Primates*, 1964, **5** (3-4), 71-80.
2. BOBBITT, R. A., JENSEN, G. D., & KUEHN, R. K. Development and application of an observational method: A pilot study of the mother-infant relationship in pigtail monkeys. *J. Genet. Psychol.*, 1964, **105**, 257-274.
3. KOGAN, K. L., & WIMBERGER, H. C. The neglected dimension of communication-methodology for immediate coding of mother-child interaction. *Amer. J. Orthopsychiat.*, 1966, in press (Abstract).
4. KUEHN, R. K., & BEACH, F. A. Quantitative measurement of sexual receptivity in female rats. *Behaviour*, 1963, **21**, 282-289.

Regional Primate Research Center
University of Washington
Seattle, Washington 98105

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 89-97.

MEMBER PERSONALITY AND GROUP COUNSELING INTERACTION*

Department of Education, Arlington State College

EMMET EUGENE KELLY[1]

A. PURPOSE

The purpose of the study was to determine the relationship between certain individual personality characteristics brought to the group and the manner in which they became visible early and late in a counseling-group setting. Measures of extroversion have been found to be positively related to total verbal activity rate in problem-solving groups (2, 11). Extroversion is indirectly related in a positive manner to task activity in groups in which an individual gives suggestions, strong opinions, and orientations, and asks questions (10). Research on problem-solving groups indicates that the more intelligent or better adjusted the individual, the more likely he is to interact in a positive social-emotional fashion rather than a negative one (3, 11). Borgatta (4) recently found that persons low in assertiveness reacted less negatively in a task group than did persons high in assertiveness.

B. SUBJECTS

Subjects of the study were 28 members of the 1963-64 Counseling and Guidance Institute of The University of Texas. Institute participants were practicing school counselors who were selected from throughout the United States to attend a nine-month advanced program, financed by the National Defense Education Act and designed to improve the counselors' professional skills and abilities.

One of the basic features of the Counseling and Guidance Institute was the Intensive Small Group (ISG) program. The Intensive Small Group experiences were designed to provide opportunities for free discussion, an individualized approach to improvement of weaknesses and extension of strengths, and a basically therapeutic situation (7). The existence of such groups provided a rich source of material for the study of interactional dynamics.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on March 31, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] Formerly at The University of Texas.

The members for the three Intensive Small Groups were selected on the basis of whom they wanted most to be with in a counseling group. The closest possible fits were made in all cases. Each of the three groups, two of nine members and one of 10, met two hours a week for a total of 26 meetings. The counselors, two men and one woman, were members of the staff of the Counseling and Guidance Institute of The University of Texas. Two held rank of lecturer in the Department of Educational Psychology and one was an assistant professor. All three persons had served partial or full internships at the Testing and Counseling Center of the University, the average internship experience having been approximately four semesters in length. The theoretical counseling orientation of the three group leaders might best be described as client-centered in its emphasis.

C. Dependent Variables

The major elements of verbal interaction analyzed were the instrumental behavior, the social-emotional behavior, and the total interactive frequency patterns of three counseling groups. Task or instrumental behavior is defined as verbal interaction which aims at the completion of group or individual tasks by asking for or giving suggestions, opinions, and orientations (1). Social-emotional behavior is defined as verbal interaction directed primarily toward establishing relationship through showing solidarity, agreement, and tension release (positive); and disagreement, rejection, and antagonism (negative). Total interactive frequency patterns are defined as interactions over time, which give an interaction profile of the individual or the group. The conceptual scheme sees social interaction as a mediating process between certain basic aspects of one's personality on the one hand and role expectations, cultural and environmental influences, on the other (1).

D. Independent Variables

Personality is conceived in this study as the personal qualities each person brings to the group. These personal qualities were further limited to certain extroversive-introversive factors as measured by the sociability ("S") scale on the Guilford-Zimmerman Temperament Survey (GZTS) and the manic ("Ma") scale of the Minnesota Multiphasic Personality Inventory (MMPI).

Persons who score high on the sociability ("S") scale of the GZTS are described as being at ease with others, enjoying their company, and being able to establish rapport readily. Persons who score low on the "S" scale of the GZTS are characterized as withdrawn, reserved, and hard to get to know (8).

Persons who score high on the manic ("Ma") scale of the MMPI are described as overactive, easily given to emotional excitement and flight of ideas (5). Persons who score low on the "Ma" scale of the MMPI are viewed as dependent, lacking drive, or lacking the ability to express drive (6).

E. HYPOTHESIS

It was hypothesized that persons characterized as extrovertive would talk more in a total way, give more opinions, ask fewer questions, and express fewer positive feelings in early group-counseling sessions than in late ones, and that the reverse would be true of persons characterized as introvertive.

F. PROCEDURE

One of the most important procedures was that of identifying the subjects and then rating them in terms of their verbal behavior in each of the three counseling groups. The two groups operationally defined were the high scorers and the low scorers on the "S" scale of the GZTS and the "Ma" scale of the MMPI.

The *high scorers* were made up of the three highest scoring persons on the "Ma" scale of the MMPI from each of the three counseling groups, and the three highest scoring persons on the "S" scale of the GZTS from each of the three groups. Since the high scorers did not necessarily score high on both the "Ma" and the "S" scale, there were 12 subjects. Six of the 12 subjects scored high on both the "Ma" and the "S" scales.

The *low scoring* subjects were made up of the three lowest scoring persons on the "Ma" scale of the MMPI from each of the three counseling groups, and the three lowest scoring persons on the "S" scale of the GZTS from each of the three groups. As was true for the high scorers, the low scorers did not necessarily score low on both scales. However six of the 12 subjects did score low on both the "Ma" and "S" scales.

Examination of *t* tests between the means of scores made by the high scorers and the low scorers by counseling groups revealed no significant differences. Although some of the *t* values were large, no difference between the groups was significant because of the small number of subjects per group.

The rationale for such a grouping procedure may be explained as follows. One of the overall purposes of the research was to discover if persons with certain personality characteristics might behave in a way relative to the personality matrix of the rest of the group. The study attempts to indicate that persons that score high or low relative to the group they are in will behave in a certain way in that group. While the mean scores of the high scorers in one group did not necessarily correspond to the mean high scores of another group, it was proposed that the relative position held in the particular group's constellation was the most important consideration in examining group interaction. Thus, absolute high scores in comparison with other population means were not used.

The design called for an examination of electronically recorded and typed verbal-interaction material of early and late meetings from each of the three ISGs. Meeting Number One was chosen as the early meeting for each of the three counseling groups. Meeting Number 21 was selected as representative of late group-counseling sessions.

After the six tapes had been converted to the typescript, a panel of three raters classified the verbal interaction. In addition to the written instructions, the raters also listened to brief portions of group-counseling meetings in order to understand better the culture of the group and the context of the interaction. Also, the investigator discussed with the raters the procedures for scoring as presented in the appendices of Bales' book on Interaction Process Analysis (1).

The unit to be classified was defined as a single verbal interaction that contained an expressed or implied subject and predicate. Each unit was first classified according to its speaker, so that frequency of interaction could be obtained. Next, the same unit of behavior was classified within one of the four categories of interaction—the positive and the negative social-emotional responses, and the task-giving and the task-asking responses.

When a unit was "borderline" between task and expressive interaction, the scorers were instructed to move the unit from task activity to expressive activity. It was assumed that a great deal of nonverbal expressive activity would be lost in utilizing only typed verbal material; thus a compensating effort was made in the direction of gaining expressive activity.

Following Bales' suggestion concerning the assessment of rater reliability, the investigator used the chi-square statistic as an index of goodness of fit. The chi-square value for the three raters exceeded a probability of .50 and was therefore acceptable, in accordance with the scheme observed by Bales (1, p. 103).

A primary statistical task, then, was to discover whether the frequency of certain verbal behavior by group members differed significantly from early to late group-counseling meetings. The single classification analysis of variance technique was used and the differences between the mean interaction counts were computed for high scorers from early to late meetings and were not found to be significant in any categories, including total interaction, for either high or low scorers.

In an effort to gain a full understanding of the differences in the mean scores as they might appear in the configurations for either the "Ma" scale, or the "S" scale, or both, multiple linear regression analysis was employed. Jennings and Whiteside (9) state that the regression technique allows for a comparison of the predictive efficiency of a restricted set of predictors with the efficiency of the full arbitrary set. The F statistic is used to determine whether or not predicting with the restricted model results in any statistically significant loss in efficiency.

In order to take full advantage of configuration differences described by the full linear regression models, the four hypotheses were further tested for expected differences between the high scores, early and late, and between the low scores, early and late. Bales' four broad categories were the criterion variables. The predictor variables, six in number for each of the two full models, were combinations of attributes (expressed as binary variables) on the "Ma" scale and the "S" scale of high scorers and low scorers during early and late group-counseling meetings.

Through the use of the Control Data Corporation 1604 Computer, the predictor variables of the two full multiple regression models were correlated simultaneously with Bales' category interaction counts of the high and low scorers for early and late group-counseling meetings. In order to determine the strength of the relationship between the predicted and criterion variables, the full multiple regression models were used to obtain coefficients of multiple determination (R^2). To determine the proportion of variance in the criterion scores that was accounted for by relationships, R^2's were computed for restricted multiple regression models derived by placing appropriate mathematical restrictions on the full models. The restricted model R^2's were then compared with the R^2's of the full models to determine whether significant differences existed; F ratios were computed to test the differences for statistical significance. The R^2 differences are reported in Table 1, along with early and late mean scores by high and low scorers on the "Ma" scale, the "S" scale, and a combination of the two.

G. Results

With regard to total interaction, the significant decrease in the coefficient of multiple determination, R^2 ($F = 7.54$), indicated significant differences ($p < .05$) in the early and late behavior of the high scorers. These differences were not always in the same direction for the different combinations of

TABLE 1

COMPARISON OF EARLY AND LATE MEAN INTERACTION COUNTS BY HIGH AND LOW SCORERS ON THE "MA" SCALE AND "S" SCALE BY BALES' FOUR CATEGORIES

Bales' categories	N	Scale	High scorers $N = 12$ Mean interaction counts Early	High scorers $N = 12$ Mean interaction counts Late	R_f^2†, R_r^2‡ F Ratio	Low scorers $N = 12$ Mean interaction counts Early	Low scorers $N = 12$ Mean interaction counts Late	R_f^2†, R_r^2‡ F Ratio
Asking questions	3	"Ma"	9.00	2.33	$R_f^2 = .3736$	7.67	6.67	$R_f^2 = .0942$
	3	"S"	6.67	5.67	$R_r^2 = .1957$	8.67	7.33	$R_r^2 = .0787$
	6	"Ma" & "S"	6.00	12.33	$F = 5.11$*	2.50	5.00	$F = .31$
Giving opinions, evaluations	3	"Ma"	14.33	10.00	$R_f^2 = .4043$	48.67	51.33	$R_f^2 = .1566$
	3	"S"	65.00	66.33	$R_r^2 = .2144$	32.00	28.33	$R_r^2 = .1367$
	6	"Ma" & "S"	38.33	105.00	$F = 5.73$*	17.50	29.33	$F = .42$
Positive social-emotional	3	"Ma"	3.33	3.00	$R_f^2 = .3192$	13.33	8.67	$R_f^2 = .2390$
	3	"S"	19.33	15.67	$R_r^2 = .2187$	13.67	6.00	$R_r^2 = .2386$
	6	"Ma" & "S"	8.33	18.67	$F = 2.65$	3.17	2.67	$F = .01$
Negative social-emotional	3	"Ma"	3.67	1.33	$R_f^2 = .3868$	4.33	6.33	$R_f^2 = .1261$
	3	"S"	8.00	6.33	$R_r^2 = .1240$	4.67	4.00	$R_r^2 = .0498$
	6	"Ma" & "S"	6.67	14.67	$F = 7.71$*	1.33	5.00	$F = 1.57$
TOTAL	3	"Ma"	30.33	16.67	$R_f^2 = .4591$	74.00	73.00	$R_f^2 = .1551$
	3	"S"	99.00	94.00	$R_r^2 = .2324$	59.00	45.67	$R_r^2 = .1370$
	6	"Ma" & "S"	56.00	150.67	$F = 7.54$*	24.50	42.00	$F = .39$

* Statistically significant at the .05 level of confidence.

† R_f^2 = coefficients of multiple determination for full linear-regression model.

‡ R_r^2 = coefficients of multiple determination for restricted linear-regression model.

scorers: while persons who scored high on both the "Ma" scale and the "S" scale interacted more frequently in late meetings than in early meetings, high scorers on the "Ma" scale alone or the "S" scale alone (but not both) interacted more frequently in the early meetings than in the late ones.

There was no significant decrease in R^2 ($F = .39$) for the early and late scorers. Consequently, this portion of the hypothesis was not statistically supported. Nevertheless, the frequency count for persons who scored low on both the "Ma" scale and the "S" scale did increase slightly from the early to late meetings as postulated. On the other hand, low scorers on the "Ma" scale alone or the "S" scale alone interacted somewhat more frequently during early than late meetings.

Again, in task-giving interactions, the decrease in R^2 ($F = 5.73$) indicated significant ($p < .05$) differences in the early and late behavior of the high scorers, though not in the direction hypothesized. An examination of the trend by high scorers on both the "S" scale and the "Ma" scale indicated a marked increase in the interaction in this category from early to late group-counseling meetings. Slight support was given the hypothesis by the decrease in the interaction count from early to late meetings by persons scoring high on the "Ma" scale only.

The hypothesis concerning low scorers was not statistically supported, as there was no significant decrease in R^2 ($F = .42$). Only in the configuration of low scores on both the "Ma" scale and the "S" scale was a trend evident, and this in support of the hypothesis that low scorers would give more opinions and evaluations in late group meetings than in early meetings.

High scorers demonstrated a statistically significant difference ($p < .05$) in the number of times they asked questions in early and late counseling-group sessions, as may be seen by the decrease in R^2 ($F = 5.11$). When trends were examined for the different constellations of high scorers, it was apparent that the high scorers on the "Ma" scale asked considerably more questions early than late. However, high scorers on both the "Ma" and "S" scales asked fewer questions during early meetings than during late group-counseling meetings.

Low scorers exhibited little difference in the frequency with which they asked questions from early to late meetings. A slight frequency increase from early to late meetings was apparent only for persons who scored low on both the "Ma" and the "S" scale.

There was no significant decrease in R^2 for either high scorers or low scorers in this positive social-emotional category; thus the hypothesis was not supported. As may be seen in Table 1, however, the differences in mean interaction counts from early to late meetings tended to be greater for the high scoring subjects than for the low scoring subjects. Of all three configurations of low scores, the configuration for low scorers on the "S" scale revealed the clearest indication of a trend in support of the hypothesis.

If the hypothesis is made that high scorers make more negative statements late than early in counseling groups, and that the reverse is true for low scorers, statistically significant support ($p < .05$) could be given this position for the high scorers. On the other hand, low scorers showed no statistically significant changes from early to late meetings.

H. DISCUSSION

Thus, in four of the five categories of interaction that were examined for high scorers in early and late group-counseling sessions, statistically significant changes ($p < .05$) occurred. Such changes did not occur, on the other hand, for any of the five interaction categories for low scorers in early and late sessions.

The significant differences that did occur in the frequency of interaction of high scorers from early to late meetings were not always in the same direction for the three high scoring subgroups. As may be seen in Table 1, the subgroup with high scorers on both the "Ma" and the "S" scales appeared to have contributed primarily to the statistically significant differences in early and late behavior.

While both the "Ma" and the "S" scale had been used previously in examining the extroversion-introversion factor, the present study found some striking differences in the interaction patterns of the three subgroups. The "manic-extroverts," persons high on both scales, might be described as continuously active in both the instrumental and expressive areas—asking questions, giving opinions and evaluations, and giving more negative reactions late than early.

It is interesting to note that subjects who scored high only on the "Ma" scale talked considerably less on the total interaction count than the other two high-scoring groups, while the subjects who scored low only on the "Ma" scale talked considerably more on the total interaction count than the other two low-scoring subgroups.

As a subgroup, the high scorers on the "Ma" scale exhibited interaction trends as predicted by three of the four original hypotheses: i.e., from early to late they asked fewer questions, gave fewer opinions, and talked less in their total interaction. In contrast, the subgroup of the low scorers on the "Ma" scale changed very little in their frequency of interaction in the Bales' categories from early to late meetings. Neither the high scorers nor the low scorers of the "S" scale demonstrated any really discernible overall trends in their patterns of interaction from early to late meetings.

I. SUMMARY

Typescripts of verbal interaction of the first and twenty-first sessions of three counseling groups of NDEA Counseling and Guidance Institute members were examined through Bales' four broad interaction categories. With the use of multiple regression analysis, a comparison was made of early and late mean interaction counts of 12 high and 12 low scorers on the manic (MMPI) and the sociability (Guilford-Zimmerman) scales. High scorers on both scales asked more questions, gave more opinions, expressed more negative feelings, and talked more than low scorers—significantly in each category—from early to late meetings. High scorers on either the manic or the sociability scale showed no significant changes. None of the subgroups of low scorers demonstrated any changes in the categorized behavior from early to late meetings.

REFERENCES

1. BALES, R. F. Interaction Process Analysis: A Method for the Study of Small Groups. Reading, Mass.: Addison-Wesley, 1950.
2. BORGATTA, E. F. Project I. Unpublished manuscript, Harvard University, Cambridge, Massachusetts, 1953.
3. ———. Analysis of social interaction and sociometric perception. *Sociometry*, 1954, **17**, 7-31.
4. ———. A systematic study of interaction process scores, peer and self-assessments, personality and other variables. *Genet. Psychol. Monog.*, 1962, **65**, 219-291.
5. DAHLSTROM, W. G., & WELSH, G. S. An MMPI Handbook. Minneapolis: Univ. Minnesota Press, 1960.
6. DRAKE, L. E., & OETTING, E. R. An MMPI Codebook for Counselors. Minneapolis: Univ. Minnesota Press, 1959.
7. EMBREE, R. B. Contract of operations. 1963-64 NDEA Counseling and Guidance Training Institute, The University of Texas, Austin, 1963.
8. GUILFORD, J. P., & ZIMMERMAN, W. S. Manual, The Guilford-Zimmerman Temperament Survey. Beverly Hills: Sheridan, 1949.
9. JENNINGS, E., & WHITESIDE, R. Multiple linear regression techniques for analysis of variance: A working paper. Research Mem. No. 8, Laboratory of Human Behavior, The University of Texas, Austin, Oct., 1963.
10. MANN, R. D. A review of the relationships between personality and performance in small groups. *Psychol. Bull.*, 1959, **56**, 241-270.
11. SLATER, P. E. Role differentiation in small groups. *Amer. Sociol. Rev.*, 1955, **20**, 300-310.

Department of Education
Arlington State College
Arlington, Texas 76010

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 99-104.

A MODIFIED WGTA FOR INFANTS IN THEIR SECOND YEAR OF LIFE*[1]

Department of Psychology, Brown University

PAUL WEISBERG AND MAE W. SIMMONS

A. INTRODUCTION

Since Harlow and Bromer's specification (3) of the Wisconsin General Test Apparatus (WGTA) for the testing of discrimination learning in monkeys by discrete-trial methods, the WGTA has been successfully employed with many species covering a wide variety of experimental problems [*cf.* Harlow (2)]. Missing from such a comparative approach is the discriminative performance of the human infant. This paper hopes to renew interest in the study of discrimination learning of motorically mature infants—an experimental concern that goes back 25 years (4, 5, 6)—by describing a modified version of the WGTA and a training procedure appropriate for infants in their second year of life. Also presented are some performance × trial functions obtained from eight infants, 12 to 16 months of age.

B. THE APPARATUS

A drawing showing the front portion of the apparatus appears in Figure 1. Part of the bottom of the rear side was cut away and in its place was hung a heavy black cloth (not shown) that extended from and was tacked to the sides of the apparatus. Also attached to the rear side where *E* was seated was a small, one-way-vision mirror. The important features on the front side of the apparatus facing *S* were two guillotine screens and, within, an 11- by 18-inch removable tray and an overhead 25-watt, yellow Christmas light. The outer screen was opaque and the inner screen transparent; both were controlled from the rear of the apparatus by a pull string. Since both screens were on one-half-inch aluminum tracks, they could easily and quickly be removed or interchanged by lifting a narrow strip of wood hinged to the roof of the apparatus.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on March 30, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported by a PHS research grant (MH 08039-01) awarded to the senior author from the National Institute of Mental Health. We wish to thank Sisters Mary Madeleine and Mary Amadaeus and the staff of St. Joseph Nursery of the Angel Guardian Home, Brooklyn, New York, for their generous cooperation during the testing of infants.

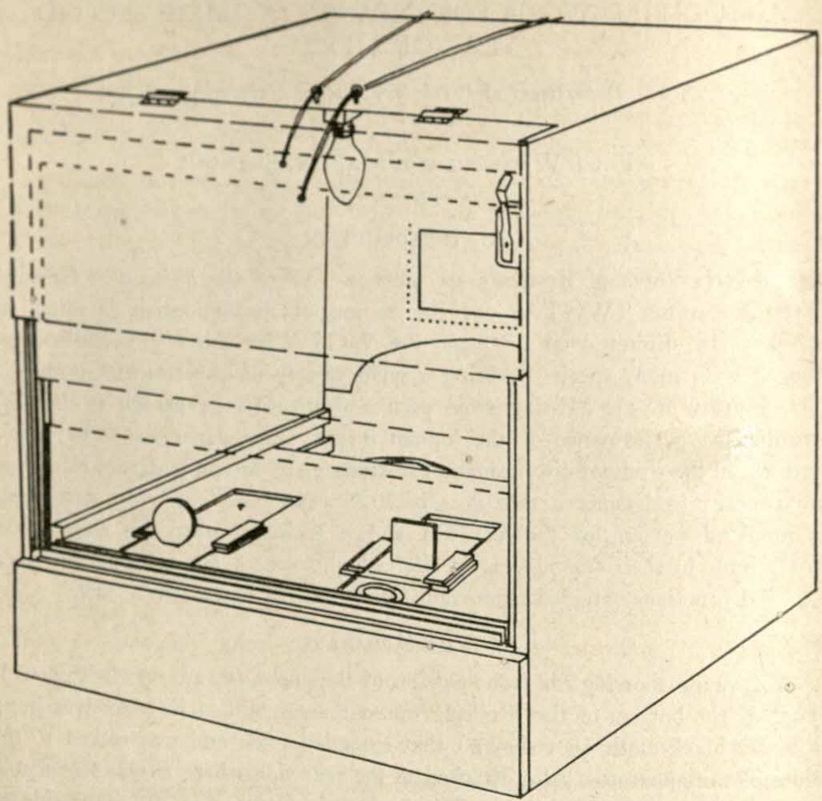


FIGURE 1
THE APPARATUS

On the tray were two recessed rectangular sections into which foodwells (one and one-half inches in diameter) had been chiseled. These foodwells (spaced eight inches apart) were covered by four-inch square grey masonite plaques which slid into place from the rear portions of the rectangular recesses. The following steps were taken to insure that the only effective response to expose entirely any one well would be to push the plaque back two inches: (*a*) pieces of wood were glued to each side of the foodwells to prevent the plaques from being lifted off the tray; (*b*) a screw was placed near the front of each well (not shown) to prevent also the removal of the plaques and to serve as a convenient starting place from which the *S* could nudge the plaque back; and (*c*) the screw at the back of the well was used as a backstop.

The discriminative stimuli, consisting of simple geometric patterns (triangle, circle, square, or inverted T) painted red, green, or yellow were mounted vertically onto the plaques. The stimuli were cut from one-fourth-inch plywood; their maximum height and width was two inches.

When the inner screen was raised halfway, a microswitch closed and started an electric timer. The clock stopped when the left or the right well was half exposed; specifically, when the sides of the plaques were moved back one inch, they knocked against a small machine screw (not shown) which in turn served as a lever actuator to two microswitches located underneath the tray.

The WGTA was constructed of one-half-inch plywood; its outer dimensions were 21 by 19 by 20 inches. When in use, it rested on a small, movable table. Counterweights were placed underneath the tray to prevent the subject from moving the apparatus; normally the WGTA weighed 22 pounds.

C. The Procedure

After an *S* familiarized himself with *E* and began to accept pieces of breakfast cereal and cookies from his hand, he was placed in a highchair and positioned in front of the apparatus. *E* sat at the rear and with the screens in a raised position began to offer *S* rewards, first by placing the reward on the front of the tray, then gradually by placing it in the foodwells. After the *S* began to pick up the food quickly, a pretraining stimulus (a two-inch porcelain pirate figure mounted onto a plaque) was placed in back of whatever well was baited. With the number of left-right baited presentations counterbalanced, the pretraining stimulus was moved gradually over one-fourth, one-half, three-quarters, and finally over the entire portion of the baited well. If an *S* during this shaping process balked at any step, he was brought back to some easier level and the process restarted. (Usually one session was needed for *S*s to learn how to uncover a closed well.) Next *E* lowered the opaque screen, baited and covered completely one of the wells with the pretraining stimulus, and raised the screen for *S* to respond. This sequence was repeated until *S* gave up his perseverative tendency to respond to the same side that had been rewarded on the previous trial and started to approach only the well covered with the pretraining stimulus. From 20 to 40 trials were needed to accomplish this.

Discrimination training was the same as the latter part of pretraining except that the pretraining stimulus was replaced by two geometric objects, one serving as the positive stimulus and the other as the negative stimulus. These stimuli were presented simultaneously and alternated from side to side according to a random series (1). Correction was discouraged by quickly lowering and raising the screens until the infant removed his hand. The intertrial interval after about 25 trials stabilized to 20 seconds. The two main criteria for ending any session were (*a*) a 35 trial per day maximum or (*b*) two consecutive failures to respond after two minutes. On the average, 17 to 29 daily trials were given to each *S*.

D. THE SUBJECTS

Five home-reared (H.R.) infants selected from an age-and-name file of parents living in the Brown Community and three orphanage-reared (O.R.) infants were tested by the procedure described above. Infants in both groups were between 12 and 16 months of age at the start of pretraining. All O.R. *S*s were tested privately in an unused nursery room which was relatively free from distracting stimuli and over a period of time in which 83 per cent of adjacent sessions were separated by 24 hours. For the remaining 17 per cent, the intersession interval was no more than 48 hours. By contrast, four H.R. *S*s were tested in their homes and 46 per cent of the adjacent sessions were separated by 24 hours, 25 per cent by 48 hours, 24 per cent from 72 to 96 hours, and 5 per cent by 120 hours. H.R. 5, however, was seen under situational and temporal conditions similar to the O.R. *S*s. The mother's presence was continuous with H.R. 1 and H.R. 4; for the other H.R. *S*s, the mother was present mostly during pretraining. Mothers and orphanage personnel were asked to omit the daily snacks the infants might normally receive. All *S*s were seen two to three hours after their last meal.

E. RESULTS

The criterion taken for evidence of learning was any consecutive run of trials (either on the same or on two successive sessions), in which there were 20 out of 25 correct responses. Figure 2A plots the performance of three H.R. *S*s who failed to satisfy this standard despite extensive training. (Trials are plotted in blocks of 30 each for the nonlearners merely for the sake of convenience; the same overall trend of chance performance would have held had the abscissa been extended with 15 trials to a block.) It is worth mentioning, however, that all of the nonlearners would have shown evidence of "learning" at least once in their records had the criterion of earlier investigators been used: e.g., Ling's criterion (4) for discrimination was 8/10 and Welch's criterion (5, 6) was 9/10.

In Figure 2B are seen the changes in choice behavior of three *S*s who solved

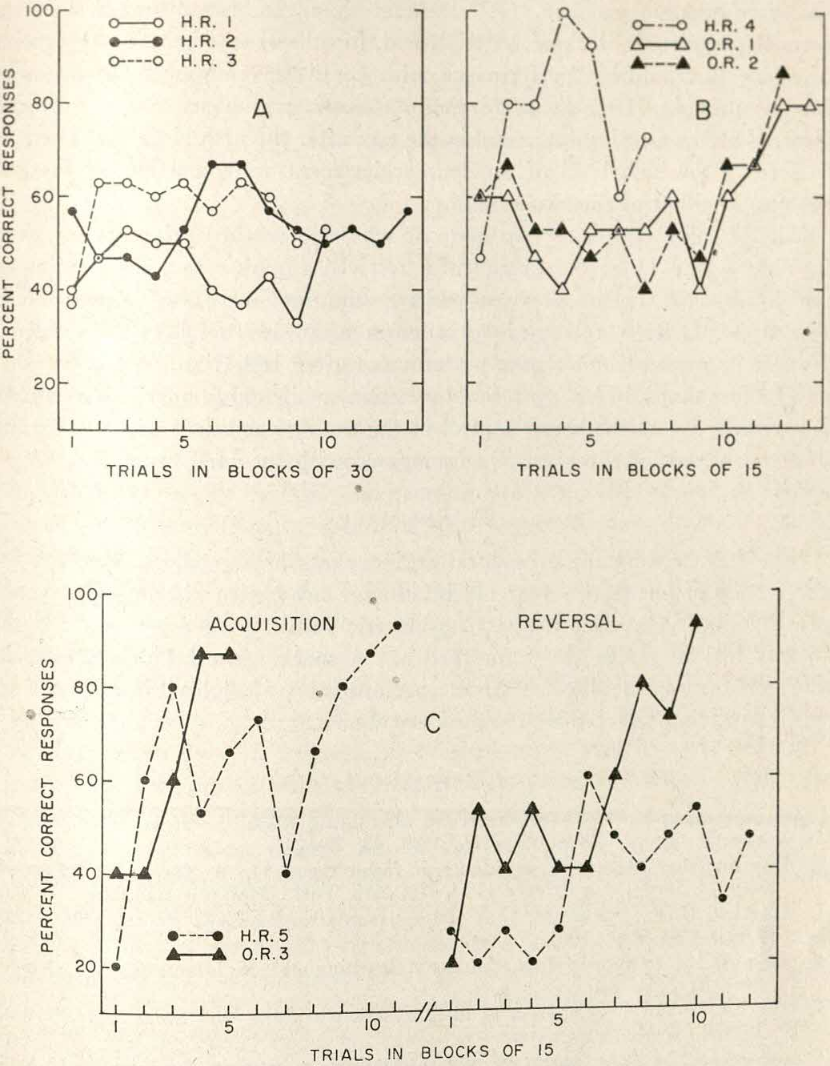


FIGURE 2

Discrimination Performance of Three Nonlearners and Five Learners

A represents the three nonlearners; B and the left section of C represent the five learners. Stimulus reversal training was subsequently given to two learners (right section of C).

the discrimination problem. (All learners were run past criterion, in some cases to complete a block of 15 trials and, in others, to determine the extent of stable discriminable performance with continued training.) The reasons for the drop in H.R. 4's performance following problem solution are not clear. H.R. 4 seemed uninterested in the task after the fifth block and attempts to recover the high level of previous achievement were met by a refusal to respond to either of the two stimuli.

The learning curves of two more *S*s who also reached criterion are given in Figure 2C. These *S*s subsequently received stimulus reversal training in which responses to the previous positive stimulus were not reinforced and *vice versa*. O.R. 3 relearned to criterion after 142 trials, while H.R. 5 did not improve beyond chance performance after 180 trials.

The fact that all O.R. *S*s reached criterion, while only two H.R. *S*s did so, is probably due to the better control in the testing conditions and the spacing of test intervals of the O.R. *S*s as compared with the H.R. *S*s.

F. Summary

The features and the procedural requirements of a modified WGTA for the testing of one-to-two year old infants on two-choice discrimination tasks are described. On each trial two geometric forms were presented and the infants had to nudge the form that had a snack reward hidden under it. Correction was not allowed. Choice performances of eight infants given extensive training under this procedure are shown.

References

1. Gellermann, L. W. Chance orders of alternating stimuli in visual discrimination experiments. *J. Genet. Psychol.,* 1933, **42**, 206-208.
2. Harlow, H. F. Learning set and error factor theory. In S. Koch (Ed.), *Psychology: A Study of a Science* (*Vol. II*). New York: McGraw-Hill, 1959.
3. Harlow, H. F., & Bromer, J. A test-apparatus for monkeys. *Psychol. Rec.,* 1938, **2**, 434-436.
4. Ling, B. C. Form discrimination as a learning cue in infants. *Comp. Psychol. Monog.,* 1941, **17**, No. 2.
5. Welch, L. The development of discrimination of form and area. *J. of Psychol.,* 1939, **1**, 37-54.
6. ———. The development of size discrimination between the ages of 12 to 40 months. *J. Genet. Psychol.,* 1939, **55**, 243-268.

Department of Psychology
Walter S. Hunter Laboratory
Brown University
Providence 12, Rhode Island

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 105-115.

VISUAL SIGNAL DETECTION, AS A FUNCTION OF AGE, INPUT RATE, AND SIGNAL FREQUENCY*[1]

Department of Psychiatry, Harvard University Medical School,
Massachusetts General Hospital

GEORGE A. TALLAND

A. INTRODUCTION

In recent years sustained attention has been studied most systematically by tests of vigilance, in which both signal frequency and the signal-to-event ratio are typically low. Signal detection in those tasks tends to drop off with continued performance (8), especially with a high rate of noise (9). This decrement has been shown to expand with age in one study (11) but not in others (1, 2). The evidence for any aging effect on vigilance is conflicting and depends largely on the experimental procedure. In as far as one can summarize the findings of several diverse experiments, it would seem that performance holds up well with age but is apt to deteriorate over 60 years if it imposes a rapid pace and a load of short term retention (4, 6, 12).

There is no conclusive evidence that a fast pace *per se*, with or without a high signal frequency, increases the difficulty of detection for older adults more than for the younger. Kirchner (6) and Welford (13) attribute the deterioration with age in performance of paced sequential tasks to failures in short term memory. It remains to be determined whether performance in such tasks deteriorates with age even though it does not involve short term memory and, if such aging effects are found, the extent to which they depend on the rates of event change and of signal frequency. A further point of interest is whether age related changes in sustained attention are progressive over the span of adulthood or occur at one or more critical stages.

The purpose of the present experiments was to investigate progressive aging effects over a span of five decades on signal detection that involves continuous visual scanning. The principal experimental variables were the rate of event change, the sequential order of rates, and the signal-to-noise ratio. Three experiments were conducted that embedded the signal amid variable noise on

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 4, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This investigation was carried out in conjunction with the Normative Aging Research of the Boston VA Clinic, and was supported by Research Grant HD 00340 and by a Research Career Award HD 15.418 from the National Institutes of Health.

a randomized temporal schedule and dictated a brisk pace, but involved no memory load, and required neither choice nor skilled operations for response.

B. EXPERIMENT I

1. *Apparatus*

Visual displays were presented by means of nine Nixie tubes, one in the center and eight spaced at equal intervals in a circle with a radius of 8.5 inches, on a four-foot square upright black panel. These vacuum tubes show any one digit between zero and nine, lighting up and turning off instantaneously. A display consisted of a variable number of digits in any combination of the nine positions. Its pattern, content, and duration were controlled by two independent mechanisms.

The content of the displays was determined by two stepping switches, one controlling five, the other four Nixies. *Part-displays* of zero, one, two, three, and four positions occurred with about equal frequency, and those of five positions half as often, within a cycle of 52 steps. Approximately one-fifth of the *part-displays* of each size contained the numeral four, so that it appeared about the same number of times in every position.

The duration of *part-displays* was controlled by a two-channel magnetic tape. In all but five out of every 60 steps the two steppers moved asynchronously, so that one part of the total display remained in view while the other part changed. At the standard rate of presentation *part-displays* were in view for 1, 2, or 3 seconds, each duration occurring with equal frequency over any one minute, and thus giving a mean *total* display of 1.1 seconds. Signals—i.e., displays that contained the numeral four, were of the same three lengths. Intersignal intervals ranged from 1 to 12 seconds, with a mean of 5.4 seconds. At the fast rate events changed at twice the standard rate. Displays containing the numeral four were recorded on two channels, responses on a third channel of a pen recorder.

2. *Subjects*

Five groups of 40 men represented each decade between 20 and 69 years. They were all employed in or had recently retired from clerical or professional jobs, and had received high school but no college education. *S*s with a record of chronic disease, more particularly of a psychiatric or neurological type, and those with uncorrected visual defects were excluded.

3. *Procedure*

*S*s were tested individually, seated in a dark room at about four feet from the panel, their eyes level with its center. In front of *S* was a telegraph key on a bench, and his instruction was to watch the changing displays and press the key each time he saw the numeral four flash in any one of the nine positions. The appearance of a four in one position while a four was showing in another counted as a signal. The test lasted for a continuous 48 minutes, displays changing at the standard rate for the initial and final 12-minute periods and at the fast rate during the middle 24 minutes.

4. *Results*

Performance was scored for errors of omission and errors of inclusion, separately over each phase of the test. Responses made within .5 second from the offset of the signal were accepted as correct. Table 1 presents correct detection for each age group at the two rates of presentation; mean ratio of individual correct scores between fast and standard schedules; difference in per cent correct between first and last period; and mean error by incorrect response. Except for the differential score, the two periods of presentation at the standard rate were pooled for the purpose of tabulation. Figure 1 shows mean correct detection as a function of age, separately for the short, intermediate, and long signal durations, at the fast and standard rates.

Analysis of variance calculated on scores of correct response showed a significant between-age group effect ($F = 3.97$, *df* 4,199), but also a significant

TABLE 1

Experiment I: Mean Number of Correct Responses at Fast and Standard Rate, Mean Difference Between Initial and Final Period at the Standard Rate, Mean Ratio Between *S*'s Correct Response at Fast and Standard Rate, and Mean Inclusion Error, at Five Age Levels

		Correct responses (per cent score)				Incorrect responses: frequency per 24 min.	
Age	*N*	Fast	Standard	Difference between first & final period (standard rate)	Ratio between fast & standard	Fast	Standard
20-29	40	55	75	—4	.75	85	29
30-39	40	53	77	—1	.71	89	29
40-49	40	51	75	—2	.69	73	35
50-59	40	50	76	0	.68	78	37
60-69	40	45	71	1	.66	84	40

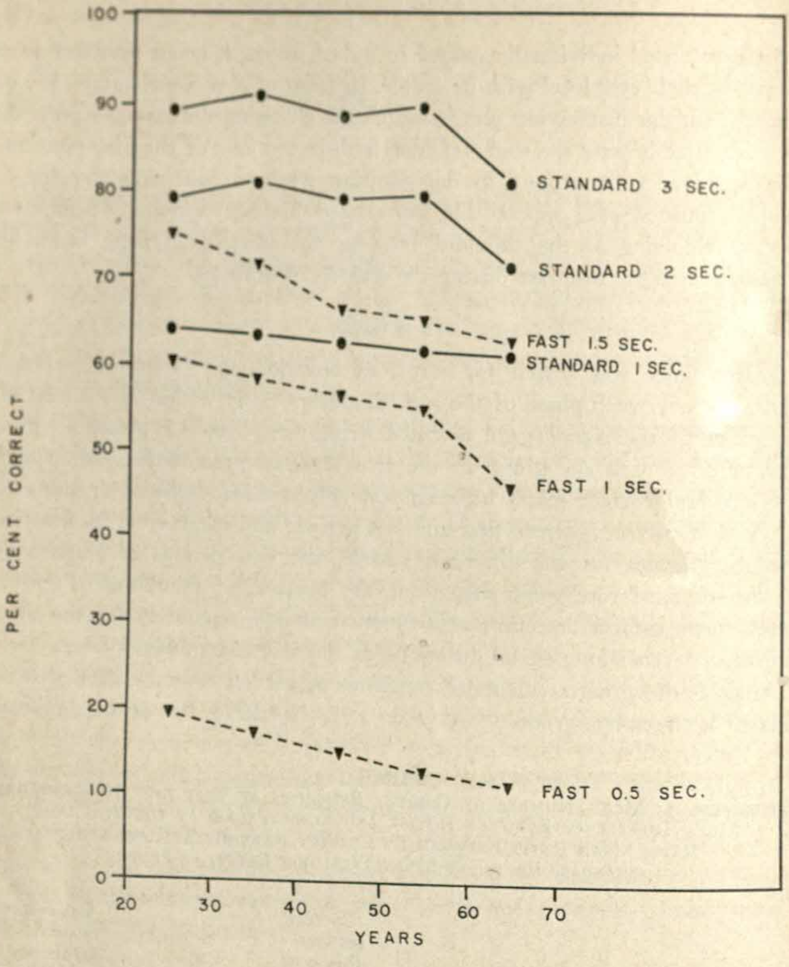


FIGURE 1
MEAN PER CENT CORRECT RESPONSE TO SIGNALS OF SHORT, INTERMEDIATE, AND LONG DURATION AT THE STANDARD AND FAST RATES, AT FIVE AGE LEVELS

interaction between age and rate of presentation ($F = 10.6$, *df* 4,191). Calculated separately for performance on fast and standard rate, correct responses under the former showed a significant between-age group effect ($F = 7.54$, *df* 4,195) and linear trend. Accuracy deteriorated stepwise over age, but none of the decades below 60 years made significantly more errors by omission than the next youngest. The combined 20's and 30's, however, performed the task significantly better than the combined 40's and 50's ($t = 2.88$) and the further drop in accuracy above the age of 60 was also significant ($t = 3.88$).

Performance at the standard rate clearly did not change progressively with age; neither the between-group variance, nor the difference between the two younger and three older decades reached statistical significance. Correct responses above the age of 60 years, however, were significantly below that of the next younger decade. Both the absolute and relative decrement in signal detection between the standard and fast rates increased with age. Measured in proportionate terms, this effect was markedly progressive and statistically significant ($F = 8.11$, *df* 4,195).

Correct responses increased from the first to the final phase at all ages except the oldest decade. This trend was most pronounced with the youngest group and least evident in the sixth decade, but between these the trend was not progressive. There were men at all ages whose accuracy in the final period fell below their initial level, but under the age of 50 they amounted to only one-third of the sample; in the sixth decade they numbered 19, and in the oldest group 22. The reversal of the effect above the age of 50 was significant both in terms of mean differences ($t = 4.79$) and in the number of *S*s whose performance deteriorated (*chi square* = 7.32). Errors of inclusion increased with age at the standard, though not at the fast rate, and also from the initial to the final phase of the test, but none of these differences was statistically significant. A unitary measure of detectability based on decision models could not be calculated because the nonsignal displays varied randomly in length and were not recorded, and, furthermore, responses made during some of these displays were treated as correct if made within .5 second from the last signal display. The effect of strategy on performance was tested by analysis of variance on combined omission and inclusion errors. Neither at the fast, nor at the standard rate did type of error interact significantly with age.

The data were also analyzed by allowing a full second for response after the offset of the signal. When performance was judged by this lenient criterion, correct responses rose by 11 per cent at the fast rate and by three per cent at the standard rate above the values shown in Table 1, about alike at all ages.

C. EXPERIMENT II

The purpose of this experiment was to test visual signal detection at a slow as well as at the standard and fast rates, and to test the sequential effect of performance rates at three ages.

1. *Method*

Subjects were drawn from the same population as in Experiment I, 12 men representing the youngest, the middle, and the oldest decade sampled. Performance was for two periods of 30 minutes separated by an hour. Alternate *S*s within an age group started with a fast-standard-fast or a slow-standard-slow schedule, and finished with the other schedule. The three phases of 10 minutes at one or another speed followed without interruption. At the slow rate displays changed half as often as at the standard rates.

2. *Results*

Mean correct and incorrect responses are listed in Table 2 for the three rates of presentation; the standard rate is shown separately within each context. There were no significant differences between the youngest and intermediate age groups, but the oldest men detected significantly fewer signals than the others at all presentation rates. All age groups performed the task more accurately at the standard rate following the fast than following the slow rate of presentation. This difference in sequential effect was significant in the youngest and intermediate age groups, and characterized 21 of the men under 50 years but only three of those above 60. With *chi square* = 13.6, the difference between the two age groups was significant.

Errors of inclusion were also fewer when the standard rate followed the fast rather than the slow rate, but in this respect no significant age difference was observed. Differences between the first and the second phase of signal detection at the slow rate pointed one way about as often as the other, at all ages alike. At the fast rate, accuracy tended to suffer rather than to improve in the final period, especially with the oldest *S*s, but none of the differences was statistically significant.

TABLE 2

EXPERIMENT II: MEAN NUMBER OF CORRECT RESPONSES AT THREE RATES OF PRESENTATION AND THREE AGE LEVELS, EXPRESSED AS PER CENT OF ALL SIGNALS, AND MEAN FREQUENCY OF INCLUSION ERRORS

		Correct responses: per cent				Incorrect responses: frequency per 12 min.			
Sequence		Slow-Standard-Slow		Fast-Standard-Fast		Slow-Standard-Slow		Fast-Standard-Fast	
Rate		Slow	Standard	Standard	Fast	Slow	Standard	Standard	Fast
Age	*N*								
20-29	12	97	77	88	54	7	15	18	44
40-49	12	91	74	86	53	6	17	16	47
60-69	12	88	68	73	46	8	21	17	45

D. EXPERIMENT III

The purpose of this experiment was to sort out the effects of fast rate and high signal frequency on signal detection at three age levels.

1. *Method*

Subjects were drawn from the same population as in Experiment I, 12 men representing the youngest, the middle, and the oldest decade sampled. Performance was continuous for 36 minutes, divided into three schedules of 12 minutes: fast, standard, and standard rate with increased frequency (16 in 60) of displays containing the numeral four. The order of the three schedules was counterbalanced within each age group.

2. *Results*

Performance evaluated by mean correct response is shown in Table 3. Significant aging effects on performance were shown only by the men over 60 years. Raising the proportion of signal events by one-third at the standard rate resulted in a 21 to 22 per cent drop in detection at all ages; false positive responses increased to a larger extent. No systematic effects attributable to the sequence of the three schedules could be determined by analysis of variance, using interaction between current and previous presentation rates as the error term. Improvement in accuracy at the standard rate following the fast schedule, however, was significantly greater than following the standard rate with increased signal frequency.

E. DISCUSSION

In regard to age, the experiments confirmed the finding that continuous performance in signal detection does not deteriorate with the advancing years, at any rate below the age of 60, if events change at a relatively slow rate. At

TABLE 3
EXPERIMENT III: MEAN NUMBER OF CORRECT RESPONSES AT FAST AND STANDARD RATES AND AT STANDARD RATE WITH INCREASED SIGNAL FREQUENCY, AND MEAN INCLUSION ERRORS AT THREE AGE LEVELS

		Correct responses: per cent			Incorrect responses: frequency per 12 min.		
Age	*N*	Fast	Standard	Standard: Increased signal frequency	Fast	Standard	Increased signal frequency
20-29	12	55	80	63	46	19	24
40-49	12	54	78	61	42	18	27
60-69	12	48	69	54	65	20	25

a faster rate accuracy deteriorates progressively with each successive decade, and there is a marked drop at about 40 and again at about 60 years. This effect is present even though the task does not demand the retention of information from previous displays. Age-related decrements in paced performance are therefore not primarily attributable to difficulties in short term retention.

The difficulties that increase with age arise from the mere requirement to scan the continuously changing displays, and to select and indicate those that have a signal property. Diminished efficiency in one or more of several functions could account for this deterioration in detection performance. Aging certainly impairs the speed of responses both in the receptor and effector mechanisms. It is unlikely, though, that oculomotor function would slow down sufficiently to affect visual scanning over an arc of 20° to 22° at the rate of .5 to 1.5 seconds, nor does the lengthening of reaction time account for the omission of responses within limits of .5 second or more. In fact, decrement with age was least marked in the detection of the shortest signals at the fast and the standard rate alike, although these allowed the least time for inspection and for response.

There is no evidence for attributing the aging effects observed to a failure in responding to signals during the short rather than during the long intervals. Those effects were unchanged by doubling the delay allowed for response. Signals were in general more likely to be detected the longer their exposure and the longer the time allowed for response, but these factors interacted with the rate of event change. Aging effects were clearly marked with signals of 1-second duration at the fast rate but barely evident at the standard rate, and therefore can hardly be attributed to an insufficient exposure of the display. Although most pronounced with older *S*s, this difference was present at all ages, and could be attributed to one or both of two conditions. First, at the standard rate a high proportion of 1-second signals remained in view without any change in the display, while at the fast rate part of the display changed every .5 second, and on occasion the change amounted to the appearance of another signal. In other words, the *S* had a full second to register the signal but not without distraction or the demand of a new decision. A second difference between the two schedules was in the frequency of responses but, even allowing for inclusion errors, the total response at the fast rate never reached twice its value at the standard rate, and dropped progressively from that ratio with increasing age. Response interference therefore offers no explanation for the age-related decline in signal detection.

The fast rate of display change was clearly the decisive factor in decreased accuracy with advancing age, and it exerted its effect principally by requiring more rapid and more frequent decisions than did the standard rate. Older men needed more time to reach the decision demanded by the task than did younger men and, while engaged in that process, may also have missed signals.

Although the response component of the task could hardly improve with practice, there is evidence from unpaced visual search experiments (10) that the operations of scanning and selection allow for learning effects. Such effects could account for the improvement of performance between the initial and final phases of Experiment I, and their likely decline with age would explain the reversal of that trend in the oldest age group. While practice over 12 minutes could reasonably result in learning effects, the process cannot go on indefinitely. Recent unpublished experiments by the author, using the same procedure as in the present study, have shown that performance does not deteriorate even over periods of two hours. It seems, therefore, improbable that attention in the present experiment slackened over time, but that this trend was masked by learning effects.

Another effect, however, could have counteracted some decrement over time in Experiment I, which, for tactical reasons, always placed the fast schedule between two spells at the standard rate. The final period always followed activity that proceeded at a brisker pace. More conclusive still was the influence of the preceding presentation rate on performance at the standard rate in Experiment II.

Expectancy and arousal hypotheses advanced to account for vigilance performance, if applicable in the present context, would both predict the sequential effects observed. The tendency of older men to show a decrement in performance over time, or following activity at the fast rate, however, is difficult to reconcile with expectancy effects which do not diminish with age (1, 3). Arousal effects, on the other hand, may reach their optimum level at lower intensities with advancing age. Past an optimum level, arousal is apt to become a disorganizing factor (5, 7), and the aging effect observed at the fast rate could indeed reflect a progressive lowering of the optimum level of arousal with advancing years.

The results of Experiment III are also more consistent with an arousal than with an expectancy theory, since they demonstrated that a fast rate of display change rather than a high signal frequency exerts a favorable delayed influence on signal detection. It is, however, debatable whether the concept of arousal associated with the intermittent signals presented in vigilance tests refers to the same hypothetical neural mechanism as is activated by the continuous sensory bombardment of the present experimental procedure. Even though the latter may not test the subcortical system associated with arousal, it is a matter of interest that performance improves as a delayed effect of increased load stress and omission errors, while reduced load stress and near perfect signal detection exert an adverse effect.

F. Summary

Men aged between 20 and 69 years were tested for accuracy in signal detection in a visual search task over periods of 30, 36 or 48 minutes. *S*s had to press a key whenever the digit four appeared in a display that changed continuously and showed randomly varied patterns of nine or fewer digits. Displays remained in view for one of three intervals and changed irregularly at one of three mean rates. One-fifth of the displays contained the digit four.

At the mean rates of 54 or 27 displays per minute performance did not change progressively with age up to 60 years. At a mean rate of 109 displays per minute errors of omission increased stepwise with each successive decade between the 20's and the 60's. Errors by incorrect response did not show a systematic trend with age. The aging effect observed is attributed to a slowing down of scanning and decision processes which match the incoming message with a model.

Performance tended to improve from the initial to the final phase, but this effect did not hold up with age. Increased signal frequency as well as a higher rate of event change reduced accuracy at all ages. Performance at the fast rate exerted a favorable delayed effect on signal detection at a slower rate with the younger and middle aged *S*s; performance at the slowest rate showed an adverse delayed effect. These observations are considered in relation to expectancy and arousal theories of vigilance.

References

1. Canestrari, R. E. The effects of aging on vigilance performance. Paper read at the Meetings of the Gerontology Society, Miami Beach, Florida, 1962.
2. Davies, D. R., & Griew, S. A further note on the effect of aging on auditory vigilance performance: The effect of low signal frequency. *J. Geront.*, 1963, **18**, 370-371.
3. Griew, S. Learning of statistical structure. A preliminary study in relation to age. In C. Tibbits & W. Donahue (Eds.), *Social and Psychological Aspects of Aging*. New York: Columbia Univ. Press, 1962. Pp. 764-769.
4. Griew, S., & Davies, D. R. The effects of aging on auditory vigilance performance. *J. Geront.*, 1962, **17**, 88-90.
5. Hebb, D. O. Drives and the C.N.S. *Psychol. Rev.*, 1955, **62**, 243-254.

6. Kirchner, W. K. Age differences in short-term retention of rapidly changing information. *J. Exper. Psychol.*, 1958, **55**, 352-358.

7. Lindsley, D. B. Emotion. In S. S. Stevens (Ed.), *Handbook of Experimental Psychology*. New York: Wiley, 1951. Pp. 473-516.

8. Mackworth, J. F. Performance decrement in vigilance, threshold, and high-speed perceptual motor tasks. *Can. J. Psychol.*, 1964, **18**, 209-223.

9. ————. Deterioration of signal detectability during a vigilance task as a function of background event rate. *Psychon. Sci.*, 1965, **3**, 421-422.

10. Neisser, U. Decision-time without reaction-time: Experiments in visual scanning. *Amer. J. Psychol.*, 1963, **86**, 376-385.

11. Surwillo, W. W., & Quilter, R. E. Vigilance, age, and response time. *Amer. J. Psychol.*, 1964, **77**, 614-620.

12. Thompson, L. W., Opton, E., Jr., & Cohen, L. D. Effects of age, presentation, speed, and sensory modality of performance of a "vigilance" task. *J. Geront.*, 1963, **18**, 366-369.

13. Welford, A. T. Changes in the speed of performance with age and their industrial significance. *Ergonomics*, 1962, **5**, 139-145.

Department of Psychiatry
Harvard University Medical School
Massachusetts General Hospital
Fruit Street
Boston, Massachusetts 02114

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 117-134.

THE STATUS OF PSYCHOLOGY IN THE UNIVERSITIES OF AUSTRIA AND GERMANY: 1965-1966*

Department of Psychology, Temple University

JAMES F. ADAMS

A. INTRODUCTION

Psychology in Austria and Germany continues to grow and develop. This article is the third in a continuing series of status reports (1, 3) that have attempted to keep American psychologists aware of the trends and developments within the universities of Austria and Germany.

Psychology within the United States has, over the years, become an increasingly self-contained enterprise. However, the fact that 85 per cent of the psychologists that hold doctorates live and work in the United States should not, as David (4, p. 315) has pointed out, "blind us to the contributions of those living and working in other lands. We share a common enterprise with common goals and aspirations. Because we are the colossus of psychology, perhaps we have an even greater responsibility to consider the views of others."

Individuals, such as Razram (7), have kept us well informed of the developments of Soviet psychology; however, there has not been a comparable interest with respect to the countries of our historical origins. Within the last five years, the author was able to find only fragmentary reports with respect to psychological developments within Austria and Germany. Wesley (9) has conducted a study of masters and pupils among German psychologists (similar to those done by Boring). He notes that only three Americans (James, Murray, and Thurstone) seem to have had any influence on German psychologists—and then only on the younger generation. Wesley (9, p. 257) suggests that the exclusion of outside influences may account for the fact that there is now "practically no comparative psychology taught or practiced in Germany and Austria and that about 27 departments of psychology, 125 teachers and 2500 psychology majors exist without the company of a single white rat." McCollum (5, p. 59) reports that there is an absence of concern with industrial psychology in the universities of West Germany. "There seems

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 4, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

to have been a wide gulf between the theoretical psychology of the university and the practical applications of industrial psychology."

Perhaps the most enlightening statement with respect to psychological research within the German universities (including Austria and the German-speaking parts of Switzerland) has been given by Metzger (6, p. 115) who is currently director of the psychology institute at the University of Münster. Metzger attributes a major part of the difficulties in doing psychological research to the *geisteswissenschaftliche Psychologie*[1] of Wilhelm Dilthey. Prior to the turn of the century, Dilthey took a rather extreme stand against psychological experimentation. His influence has continued up to the present time and has had an effect on the allocation of funds for, and the tolerance of, psychological research. Metzger points out further that psychology has still not found a comfortable home within the university faculties. If it is found within the science faculty of a university, it is considered not sufficiently rigorous and scientific; if it is located in the humanities, it is looked upon with distrust as being too scientific. Staff appointments present another problem as "the planned appointments are debated in commissions, in which the psychologist finds himself amidst a majority of philosophers, historians, and philologists. And these very nearly unanimously venerate Dilthey and still believe in his verdict against experimental methods." Metzger may gain some comfort from an examination of Table 1, within the present article, which shows that, in spite of the difficulties that he has enumerated, there is a phenomenal growth within psychology in his country—growth having been measured in terms of additions to facilities, faculty, and students. The direction and the quality of research are, of course, quite a different problem.

In the last status article (3) it was reported that a monumental publishing effort had been begun in Germany. The 12-volume *Handbook of Psychology,* under the coordinating editorship of Professors K. Gottschaldt, Ph. Lersch, F. Sander, and H. Thomae, is almost two-thirds completed. To date, seven volumes have appeared and two more will be released in the near future. Among the volumes now available are those covering the topics of developmental psychology, personality, educational psychology, psychological testing, learning and thinking, and motivation. Professor Thomae states in a personal communication to the author "as we try to give a fairly representative report on problems and findings of the international present day psychology, the enterprise might be of interest to American readers." With respect to this

---

[1] There is no English equivalent for this term, but in general it implies a psychology that restricts itself to methods of interpretation and explanation as they are found in the scientific treatment of poetry and fiction.

TABLE 1
FACULTY, STUDENTS, AND FACILITIES

Institute data	Year	Austria*			Germany**						
		1	2	3	1	2	3	4	5	6	7
Number of full-time institute faculty	1965-66	4	2	7	9	3	10	7	7	10	11
	1960-61	4	3	3	6	1	7	3	2	3	9
	1955-56	4	3	3	2	1	5	2	2	3	7
Number of part-time institute faculty	1965-66	3	3	2	10	8	7	0	8	4	7
	1960-61	3	0	1	3	8	7	1	3	3	7
	1955-56	3	0	1	7	8	8	0	3	2	4
Number of students enrolled in psychology	1965-66	240	60	320	550	1000	600	800	630	450	500
	1960-61	200	30	230	250	575	180	120	60	200	200
	1955-56	160	50	120	250	575	150	100	60	60	120
Number of students in diploma program	1965-66				270	80	250	200	70	230	370
	1960-61		Not given		200	25	120	40	35	110	120
	1955-56				120	25	117	40	35	12	70
Number of Ph.D. candidates	1965-66	21	25	90	20	17	30	25	10	13	15
	1960-61	18	9	80	10	12	30	15	12	8	15
	1955-56	15	9	60	26	12	30	15	12	6	10
Number of rooms in the institute	1965-66	7	5	32	56	52	58	41	12	31	27
	1960-61	7	5	17	44	41	18	10	5	18	25
	1955-56	5	4	17	28	41	18	10	5	17	10
Number of mechanics employed	1965-66	0	1	2	1	8	4	2	0	2	1
	1960-61	0	0	1	1	4	4	0	0	2	1
	1955-56	0	0	1	1	4	4	0	0	1	0

TABLE 1 (*continued*)

Institute data	Year	Germany**											Per cent of increase over 10 years
		8	9	10	11	12	13	14	15	16	17	18	
Number of full-time institute faculty	1965-66	3	12	6	17	8	9	14	16	6	9	4	
	1960-61	3	8	3	17	4	7	6	5	4	6	4	
	1955-56	3	2	3	9	4	4	7	5	1	6	3	120
Number of part-time institute faculty	1965-66	1	5	9	20	6	4	18	6	3	2	3	
	1960-61	1	5	2	20	3	10	4	11	3	3	2	
	1955-56	1	1	2	2	3	6	11	11	2	2	2	63
Number of students enrolled in psychology	1965-66	400	800	450	100	250	500	800	600	200	300	500	
	1960-61	100	700	120	100	150	470	650	200	100	200	250	
	1955-56	80	350	120	200	50	290	650	200	60	40	80	167
Number of students in diploma program	1965-66	180	370	380	20	60	200	550	300	150	230	200	
	1960-61	55	200	60	20	60	80	300	60	35	140	50	
	1955-56	50	55	60	20	30	42	220	60	10	25	40	299
Number of Ph.D. candidates	1965-66	48	15	6	7	12	5	25	34	6	10	15	
	1960-61	25	10	6	14	12	7	20	6	4	18	12	
	1955-56	23	5	6	1	12	4	25	6	4	9	12	49
Number of rooms in the institute	1965-66	58	29	44	40	19	44	28	64	40	40	25	
	1960-61	21	15	13	40	19	38	22	19	14	17	24	
	1955-56	21	10	13	40	6	33	14	19	11	17	24	107
Number of mechanics employed	1965-66	7	1	2	1	0	5	2	5	1	2	3	
	1960-61	3	1	1	1	0	4	2	3	0	1	2	
	1955-56	3	1	1	1	0	3	1	3	0	1	1	92

* Austria: 1 = Graz, 2 = Innsbruck, 3 = Vienna.

** Germany: 1 = Berlin, Free; 2 = Berlin, Humboldt; 3 = Bonn; 4 = Cologne; 5 = Erlangen Nürnberg; 6 = Frankfurt; 7 = Freiburg; 8 = Göttingen; 9 = Hamburg; 10 = Heidelberg; 11 = Leipzig; 12 = Mainz; 13 = Marburg; 14 = Munich; 15 = Münster; 16 = Saarbruecken; 17 = Tubingen; 18 = Würzburg.

monumental publishing effort, a comment made by Goodwin Watson (8, p. 771) is worth repeating. "German psychology is still more thoughtful, more qualitative . . . more concerned with wholes . . . more interested in achieving insight. . . ." The *Handbook of Psychology* exemplifies a continuing German tradition: i.e., the tradition of periodically synthesizing a body of knowledge. Considering the enormity of this task today, and the willingness of the German psychologists to undertake it, one is left with a feeling of awe (both for the task and for the strength of conditioned responses in education).

A number of changes have occurred in the directorships of the Psychology Institutes of Germany. Professor Gottschaldt has left Humboldt University in Berlin and is currently the director at Göttingen. Professor Klix has recently assumed Gottschaldt's duties at Humboldt University. Professor Toman has filled the vacant directorship at Erlangen-Nürnberg. Professor Rudert resigned from the position of director at Heidelberg in 1963, but continues his teaching and research. Professor Graumann has assumed the directorship at Heidelberg. Professor Witte has left Tubingen to become codirector with Professor Metzger at Münster. Professors Mittenecker and Bergius have become codirectors of the Tubingen institute.

An interesting development is occurring in the structure of psychology within the universities of Germany. In 1960 there were 21 institutes of psychology within the 21 universities of Austria and Germany that trained professional psychologists at the doctoral level. Each of these institutes had a single director (Austria continues in the same pattern). Today, a number of the West German universities have added institutes for the applied and social areas of psychology. For example, Cologne now has three institutes, each with a separate director. This has increased the number of institute directors or codirectors to a total of 28. There are also several universities where a new administrative aspect of psychology has been incorporated under the leadership of the institute director. For example, at Mainz, Professor Görres is subdirector of Applied and Depth Psychology under the leadership of Professor Wellek. The inauguration of these new institutes would seem to parallel the situation as it exists in the United States where we may have departments of psychology, social psychology, and educational psychology within the same university. In addition, it may indicate a greater concern with the applied aspects of psychology than existed before, both within the university and the community at large. It will be interesting to follow this development in the future with respect to its continuation and spread within the universities and its probable influence on the direction taken by psychological research.

In view of the increasing number of psychological institutes and, consequently, the number of directors, it might be conjectured that this trend would be accompanied by a change in the average age of institute directors. Apparently, this conjecture, on the part of the author, is erroneous. Computing the average age of institute directors for the years of 1955, 1960, and 1965 reveals that the mythical individual described by this statistic is about 53 years old and grows no older nor younger. The youngest director is Professor Klix at Humboldt University in Germany (38 years), and the senior director of all the institutes is Professor Weinhandl at the University of Graz in Austria (69 years). Professor Weinhandl has been seriously ill for well over a year and hospitalized for much of this time. In spite of this, he has conducted his duties as director of the Graz institute. Doctoral examinations, faculty and student conferences, and administrative duties have all continued while Professor Weinhandl has been convalescing from the amputation of his left leg. It would seem that a tribute to the dedication of Professor Weinhandl would not be out of order!

Adams (2) has found that American psychologists do not have access to Austrian and German research publications. In an examination of *Psychological Abstracts* he has noted that less than 30 per cent of the major research publications have been referenced. An examination of the German publications that have been abstracted since that time reveals that English readers still have little access to information concerning what is being published in German-speaking countries. This is a matter of current concern to Dr. Philip J. Siegmann (editor of *Psychological Abstracts*) and he has indicated, through personal correspondence with the author, that he is working on possible ways to increase the flow of abstracted research-publication information. Suggestions of European psychologists will be most welcome.

In the following sections, each of the university institutes of Austria and Germany is treated separately. In addition to biographical data on the directors, research interests of the respective institutes are noted. It is hoped that psychologists in the United States will not hesitate to correspond with institutes of research orientation similar to their own. Perhaps personal correspondence will facilitate the communication problems that continue to exist.

B. Austria

1. *Karl-Franzens University, Graz*

Professor Ferdinand Weinhandl, director of the psychology institute of Graz, was born January 31, 1896, in Judenburg, Steiermark (Austria).

Professor Weinhandl studied with V. Benussi, A. Meinong, and E. Martinak, obtaining his doctoral degree at Graz with a dissertation entitled "Experimentelle Untersuchungen zur Analyse des Verstehenserlebnisses Wort und Bedeutung." His present philosophical-psychological orientation he attributes to the influence of Goethe, Ehrenfels, and Gestalt psychology. The current areas of research at Graz are Gestalt psychology;[2] educational psychology; Psychagogik;[3] projective tests (especially the Gestaltlegetest of Weinhandl); characterology[4] and personality; the psychology of feeling and behavior; the psychology of space and orientation; the psychology of arts and language; social psychology; the psychology of physiognomical marks; graphology; theory of consciousness and memory research; neurovegetative processes; industrial psychology; child and geriatric studies; the psychology of expression; and studies along the lines of those by Arnheim, Heider, Metzger, Eysenck, Lienert, and Rohracher.

2. *Leopold-Franzens University, Innsbruck*

The present director of the Innsbruck institute is Professor Ivo Kohler, who was born July 27, 1915, at Schruns, Austria. His mentors were Th. Erismann, F. Mayer-Hillebrand, and R. Strohal. Professor Kohler's doctoral degree was awarded at Innsbruck, with his dissertation "Der Einfluss der Erfahrung in der optischen Wahrnehmung." The research interests of Innsbruck are concerned with perceptual investigations employing distorting lenses (similar to the work of Stratton); the obstacle sense of the blind; intelligence and personality testing; ethical behavior; and space, time, motion, color, and acoustical research.

3. *University of Vienna, Vienna*

Professor Hubert Rohracher, born April 24, 1903, at Lienz, Tyrol (Austria), is the present director of the institute. His teachers were E. Becher and Th. Erismann, and his dissertation "Die Methodenlehre, Fechners" was the basis for his Ph.D. degree awarded at Munich. The Vienna institute is engaged in research involving neuropsychological problems, learning, personality, microvibrations, personality testing, statistical methods, and accident investigation.

---

[2] Research areas are those given by the directors and faculty members as representing the interests of the respective institutes. At the discretion of the author, some areas were deleted as being too broad to indicate specific interests: e.g., "general psychology."

[3] As nearly as the author could establish, there is no English psychological equivalent for *Psychagogik*. The concept covers the areas of educational psychology, mental hygiene, and guidance.

[4] Characterology refers to individual differences—the constants of personality—that are largely uninfluenced by environmental conditions. It also refers to the area of personality best assessed by projective methods.

C. Germany

1. *Free University of Berlin, West Berlin*

Professor Hans Aebli, born August 6, 1923, at Zurich, Switzerland, shares the directorship (on an alternate-year basis) with Professor Hörmann. Professor Aebli received his Ph.D. degree at Geneva under the supervision of Piaget. He credits his present interest in cognitive growth to the influence of Piaget. Professor Aebli's dissertation was entitled "Didactique psychologique: Application à la didactique de la psychologie de Jean Piaget."

Professor Hans Hörmann was born October 23, at Ulm, Germany. Professor Hörmann studied under von Allesch and Wilde at Göttingen and received his doctorate there with the dissertation "Über die Auswertbarkeit eines schriftlichen projektiven Verfahrens." He attributes his current interest in perception to von Allesch, and his attention to learning theory and psychological diagnostics to Wilde.

The research interests of the institute pertain to perception, learning, noise and arousal, psycholinguistics, cognition and cognitive growth, social perception, methodology, and industrial psychology.

2. *Humboldt University (Formerly Friedrich-Wilhelm University), East Berlin*

Professor Friedhart Klix was born October 13, 1927, in Friedersdorf (in Saxonia). He studied under Professors Grimm, Segal, Gottschaldt, and Straub, and received his doctorate at Humboldt University. His doctoral dissertation was entitled "Die relative Sehgrösse der Gegenstände bei Eigenbewegung des Wahrnehmenden." Professor Klix indicates that he is presently interested in cognitive structures, models of human information processing, invariants common to perception, and concept formation and thinking. The research areas of the Humboldt institute encompass the application of mathematical models in psychology, cybernetics as directed to applications in engineering psychology, and the use of psychodiagnostics in clinical psychology and psychotherapy.

3. *Rheinische Friedrich-Wilhelm University, Bonn*

Professor Hans Thomae, born July 31, 1915, in Winkl, Bavaria (Germany), is director of the institute at Bonn. His preceptors were Rieffert and E. Rothacker. The title of his doctoral dissertation was "Bewusstsein und Leben: Versuch einer Systematisierung des Bewusstseinproblems," and his degree was earned at Bonn. Professor Thomae credits his present orientation to Nietzsche, E. Rothacker, Ph. Lersch, G. W. Allport, and the sociocultural approach of American personality theory.

Professor Adolf Martin Däumling, born February 12, 1917, in Munich, and formerly at the University of Munich, has been appointed to the newly established professorial chair for Applied Psychology and is codirector of the Bonn institute. In addition, he is chief of the Child Guidance Clinic. Professor Däumling studied under R. Pauli and Ph. Lersch at Munich where his doctorate was awarded upon a dissertation entitled "Experimentelles Ausdruckszeichnen." He credits his orientation to W. Stern, C. G. Jung, and Ph. Lersch.

The research interests of the Bonn institute include motivation theory (especially the theory of decision making), personality theory, longitudinal studies of aging and adult development, psychomotor development, social psychology of family life, trial and error learning, clinical psychology and psychotherapy, problems of cognitive dissonance, and replications of attitudinal studies undertaken 35-40 years ago with juveniles.

4. *University of Cologne, Cologne*

There are now three psychology institutes at the University of Cologne. Each is found in a different faculty of the University. Professor Undeutsch continues to direct the institute in the natural sciences; Professor Anger directs the Institute for Social Psychology which is found in the department of Social and Economic Sciences; and Professor Salber heads the institute located in the Humanities.

Professor Udo Undeutsch was born December 22, 1917, at Weimar, Germany. His mentor was Fr. Sander and his doctorate was taken at Jena with a dissertation entitled "Gestalttypologische Untersuchung der Sinnerfüllung optischer Komplexe." The research interests of this Cologne institute center around psychological diagnostic methodology, adolescent and child psychology, forensic psychology, secondary school selection, and traffic psychology.

Professor Hans Anger was born June 13, 1920, at Essen, Germany. He studied with H. Düker at Marburg (where his doctorate degree was received) and with L. L. Thurstone at the University of Chicago. The title of his doctoral dissertation was "Über die Wirkung von Heimdrüsenhormonen auf die geistige Leistungsfähigkeit minderbegabter Kinder." Professor Anger (who is also a physician) feels that his present philosophical-psychological orientation could best be labeled "eclecticism." The research interests of the institute are in the area of social psychology and include such topics as group interactions and processes, mass communication, and attitude development and change.

Professor W. Salber was born March 9, 1928, at Aachen, Germany. He studied under Rothacker, Sander, and Thomae at Bonn and received his degree there with a dissertation entitled "Urteil, Entschluss und Entscheidung." His present orientation is derived from morphology. The research areas of Professor Salber's institute are found in the philosophy of science, educational psychology, motivational research, and psychoanalysis.

5. *Friedrich-Alexander University, Erlangen-Nürnberg*

Professor Walter Toman, born March 15, 1920, in Vienna, Austria, is director of the Erlangen institute. Professor Toman studied under Rohracher, Kraft, v. Bertalanffy, and Aichhorn. His dissertation "Experimenteller Beitrag zum Verstehen" was written and his degree awarded at Vienna. Professor Toman acknowledges the influence of the "Wiener Kreis" and Freud on his present-day thinking. The chief research interests of the Erlangen institute have to do with the development of motivation, personality theory, psychopathology, social psychology (with particular reference to the psychological environment, early social relationships, and family constellations), perception, learning theory, and language.

The University of Erlangen has been merged with the Hochschule für Wirtschafts und Sozialwissenschaften of Nürnberg. The chief function of the psychology institute at Nürnberg is to provide psychological training for students enrolled in economics, applied sociology, and other subjects. Professor Theodor Scharmann, born on July 12, 1917, in Kreuzlingen, Switzerland, is the director of this institute. Professor Scharmann studied under Gruhle, Hellpach, Jaspers, Gelb, and Wertheimer. His doctoral dissertation was entitled "Saelde (Schieksal) in der ritterlichen Dichtung d. 12. n. 13. Jhdts." His present philosophical-psychological orientation is attributed to the influence of Gestalt and depth psychology together with cultural anthropology. The research interests of the Nürnberg institute encompass motivation theory and the theory of social action, general and social perception, attitude research, the psychological theory of economic behavior, and rehabilitation psychology.

6. *Johann Wolfgang Goethe University, Frankfurt*

There are now two psychology institutes at the University of Frankfurt. Professor Rausch has filled the position vacated by Professor Thomae. This institute is located in the Natural Science Department. Professor Süllwold has been appointed to the directorship of the new institute in the Philosophy Department. The two institutes work together in the training of students.

Professor Edwin Rausch was born February 1, 1906, at Baumholder, Germany. His major professors were M. Wertheimer and W. Metzger. His doctoral degree was taken at Frankfurt with a dissertation entitled "Über Summativität und Nichtsummativität," and he credits his present orientation to the influence of Gestalt theory. Research areas of the institute are perception, learning, thinking, memory, sensori-motor, diagnostics, the psychology of expression, motivation, and *Kategorialanalyse*.[5]

Professor Fritz Süllwold, born August 6, 1927, in Herne, Germany, is the director of the second Frankfurt institute. His major professors were J. von Allesch and K. Wilde. Professor Süllwold's doctoral degree was earned at the University of Göttingen and his dissertation topic was "Experimentelle Untersuchungen über die Rolle des Einfalls im Denkprozess." The main research areas of this institute are the psychology of thinking, factor analytic studies of intelligence, learning theory, attention, educational measurement including test construction, attitude measurement, and experimental design.

7. *Albert-Ludwig University, Freiburg, in Breisgau*

Professor Robert Heiss, born January 22, 1903, at Munich, Germany, is director of the institute at Freiburg. His teachers were N. Hartmann and K. Jaspers; his doctoral dissertation, "Über die Negation," was written and his degree awarded at Göttingen. Professor Heiss attributes his present philosophical-psychological approach to Hegel and Freud. The research interests of the Freiburg institute are in the areas of personality, projective techniques, graphology, pharmaco-psychology, psychosomatic medicine, and forensic psychology (juvenile delinquency). Research also continues on The Color Pyramid Test, a projective technique developed at Freiburg.

8. *Georg-August University, Göttingen*

Professor Kurt Gottschaldt, born April 25, 1902, in Dresden, Germany, assumed the directorship of the Göttingen institute in February of 1962. Professor Gottschaldt studied under the Gestaltists W. Köhler, M. Wertheimer, and Kurt Lewin and received his doctorate at Friedrich Wilhelm University in Berlin. His doctoral dissertation was entitled "Über den Einfluss der

---

[5] *Kategorialanalyse* is a philosophical term and refers to the classifying of psychological phenomena into philosophical categories such as thinking, perception, etc.

Erfahrung auf die Wahrnehmung von Figuren I u. II Der Aufbau des kindlichen Handelns." Professor Gottschaldt attributes his present philosophical-psychological orientation to the influence of Gestalt psychology and field theory. Research at the institute encompasses problems in the areas of personality, developmental psychology, psychopathology in children, diagnostics, and perception.

9. *University of Hamburg, Hamburg*

Professor Peter R. Hofstätter, born October 20, 1913, in Vienna, Austria, is director of the Hamburg institute. His doctoral dissertation, "Testuntersuchungen an japanischen Kindern und das Reifungsproblem," was completed at the University of Vienna under the direction of Karl and Charlotte Bühler. Professor Hofstätter attributes his present philosophical-psychological orientation to the Bühlers, E. Brunswik, and to L. L. Thurstone. The Hamburg institute is doing research in clinical psychology (test construction and counseling), social psychology (attitude scaling and group dynamics), and in experimental psychology (decision and information theory).

10. *Ruprecht-Karl University, Heidelberg*

Professor D. F. Graumann, born March 31, 1923, in Cologne, Germany, is the current director of the psychology institute. Professor Rudert resigned from the directorship in the summer of 1963, but continues his research and teaching duties. Professor Graumann studied with M. Krudewig, E. Rothacker, and H. Gruhle, and obtained his doctoral degree in Cologne with the dissertation "Die Kriterien des Einfallserlebens: eine theoretische und experiementelle Analyse." His philosophical-psychological orientation is towards the convergence of phenomenological and operational conceptions and methods in humanistic psychology. The research interests of the Heidelberg institute are the development and refinement of methods for the descriptive analyses of experience and behavior, psycholinguistics, and the social psychology of cognition.

11. *Karl-Marx University, Leipzig*

Professor Werner Fischel, born October 21, 1900, in Saarburg, Rheinland, is director of the Leipzig institute. His pedagogues were K. Marbe, R. Hertwig, and O. Koehler, and his doctoral dissertation, "Haben Vögel ein 'Zahlengedachtnis'?" was written and his degree awarded at Munich. The research interests of the institute are vocational guidance, the psychology of thinking and aptitude, educational psychology, general comparative psychology, and animal psychology.

12. *Johannes Gutenberg University, Mainz*

Professor Albert Wellek, born October 16, 1904, in Vienna, Austria, is the director of the Mainz institute. Professor Wellek's mentors were F. Krueger and E. Brunswik. His doctor's degree was received at Vienna with a dissertation entitled "Doppelempfinden und Programm-Musik." Professor Wellek credits his present orientation to F. Krueger's *Genetische Ganzheitspsychologie*. The research interests of the Mainz institute are in characterology, diagnostics, acoustics, aesthetics, the psychology of music, expression, motivation ("will"), projective techniques, graphology, cultural patterns of musical hearing, mythology, psychoclimatology (the psychic influence of the weather), social and child psychology, and hypnosis.

Since 1962, an area of applied and depth psychology has been in operation at the Mainz institute. Professor Albert Görres has been appointed as subdirector of this program.

13. *Philipps University, Marburg*

Professor Heinrich Düker, director of the Marburg institute, was born November 24, 1898, at Dassel, Germany. His teachers were G. E. Müller and N. Ach. Professor Düker's doctoral dissertation, " 'Über das Gesetz der speziellen Determination': Ein experimenteller Beitrag zur Lehre vom Willen," was written and his degree taken at Göttingen. He attributes his present philosophical-psychological orientation to "empiricism." The research interests of the institute are efficiency and effectiveness, pharmaco-psychological research (hormones and stimulants, barbiturates, vitamins, glutamic acid, and lysergic acid), statistics, perception, motivation and cognition, industrial psychology, test construction, psychophysics (galvanic skin reflex, feelings, and emotions), forensic psychology, educational psychology, and the history of psychology.

14. *Ludwig-Maximilians University, Munich*

Professor Philipp Lersch, born April 4, 1898, at Munich, Germany, is director of the institute. His professors were E. Becher, J. Geyser, and K. Huber. The title of his doctoral dissertation was "Der Traum in der deutschen Romantik" and his degree was awarded at Munich. He gives primary credit in his present thinking to the influence of M. Scheler. Research at the Munich institute encompasses personality, work and industrial psychology, the psychology of expression, diagnostics, motivation, social psychology, and experimental psychology.

15. *Westfälische Wilhelms University, Münster*

The Münster institute now has two divisions and two directors. One of these divisions is for general psychology; the other, for applied psychology. Professor Metzger is presently the director in chief of the institute, but it is planned that this position will rotate regularly between Professor Metzger and Professor Witte.

Professor Wolfgang Metzger was born July 22, 1899, at Heidelberg, Germany. He studied under M. Wertheimer, W. Köhler, Kurt Lewin, and K. Jaspers. Professor Metzger's doctoral dissertation was entitled "Über die Vorstufen der Verschmelzung von Figurenreihen, die am ruhenden Auge vorüberziehen," and his degree was awarded at Friedrich-Wilhelm University in Berlin. He attributes his present philosophical-psychological orientation to the influence of J. G. Hamann, M. Wertheimer, and W. Köhler's Berlin school of Gestalt theory.

Professor Wilhelm Witte has resigned from the directorship of the Tubingen institute and has been at Münster since April of 1964. Professor Witte was born January 8, 1915, at Hattingen, Ruhr (Germany). His professors were N. Ach, W. Hellpach, and K. Jaspers. His dissertation, "Die Metaphysik von Ludwig Klages," was written and his degree received at Heidelberg. Professor Witte attributes his present philosophical-psychological position to the influence of W. Köhler, M. Wertheimer, and K. Lewin.

The research interests of the institute cover phenomenology; perception; personality theory; theoretical, developmental, social, mathematical, educational, diagnostic, and applied psychology; the psychology of will, needs, performance, feeling, expression, and communication; psychophysics; developmental disturbances; and film psychology.

16. *University of Saarbruecken, Saar*

Professor Ernest E. Boesch, born December 26, 1916, at St-Gall, Switzerland, is director of the institute. His pedagogues were E. Claparede, J. Piaget, A. Rey, and O. Pfister; his dissertation, "L'organisation d'un service de psychologie scolaire," was completed and his degree awarded at Geneva, Switzerland. Professor Boesch acknowledges the influence of his combined experimental, clinical, and psychoanalytic training. More recently he states that his research in Thailand (1955-1958), as director of the International Institute for Child Study, has strongly guided the direction of his work and thinking. In addition to clinical child psychology, the research interests of the institute now focus mainly on problems of social and cross-cultural child psychology: i.e., social and cultural factors influencing the growth of personality, and the methodology of cross-cultural comparisons.

17. *Eberhard-Karls University, Tubingen*

Since April of 1965, the Tubingen institute has been under the codirectorship of Professors Mittenecker and Bergius.

Professor Rudolf Bergius was born in Carlshof, Germany, April 22, 1914. His mentors were Adhémar Gelb, Hans Keller, Eduard Spranger, and Hans Rupp. The doctoral dissertation of Professor Bergius was entitled "Die Ablenkung von der Arbeit durch Lärm und Musik und ihre strukturtypologischen Zusammenhänge." He attributes his present philosophical-psychological orientation to the influence of Köhler and Wertheimer in combination with the modern methods of behavioral research.

Professor Erich Mittenecker was born on June 26, 1922, at Wiener Neustadt, Austria. His professors were H. Rohracher and V. Kraft and his dissertation, "Erleben und Objektivität," dealing with epistemological problems in psychology, was written and his degree awarded at the University of Vienna. Professor Mittenecker credits the Vienna Circle of modern philosophy (Schlick, Carnap, etc.) as being influential in his present thinking.

The research interests of the Tubingen institute center around phenomenology, developmental and social psychology, learning and thinking, personality theory and diagnostic techniques, expression, mathematical methods, and psychophysics.

18. *Julius-Maximilians University, Würzburg*

Professor Wilhelm Arnold, born October 14, 1911, at Nürnberg, Germany, is director of the Würzburg institute. Professor Arnold studied under A. Fischer and R. Pauli and obtained his doctoral degree at Munich with a dissertation entitled "Über die Wahrnehmbarkeit akustischer Signale." He acknowledges the influence of J. Geyser and R. Hönigswald as being most influential in his present thinking. The research interests of the Würzburg institute are personality and personality assessment, ability, social psychology, educational psychology (especially programmed learning), forensic psychology, and professional consulting.

D. UNIVERSITY RESEARCH INTERESTS

When the research interests of the institutes are examined, Wesley's comments on the absence of animal and comparative psychology seem to be substantiated. Only Leipzig indicates an interest in this area. McCollum's (5) impressions on the activity in industrial psychology within the West German universities do not receive support from this survey. Nine of the 16 West German universities list research activities which might be considered within the industrial domain. Industrial psychology is a continuing research concern.

Looking for trends in university research between the years of 1955 (1) and 1965 reveals some rather interesting data. It appears that interest in psychological-test development is decreasing; 19 universities reported work in this area 10 years ago; 15 do so today. Research in statistics, experimental design, mathematical models, and cybernetics is on the increase; and six institutes report activity in these areas—double that of 1955. Investigations in developmental psychology have decreased by one-third (from 12 to eight universities); social psychology has remained constant (12 universities); and general concern with man (personality, consciousness, phenomenology, feeling, expression, cognition, motivation, etc.) is still the major research concern (18 universities). The traditional areas (learning, time, space, memory, psychophysics, music, etc.) continue to receive attention (15 universities); and the number of universities involved in clinical research remains constant at seven.

E. Faculty, Students, and Facilities

In Table 1 can be seen the size of the faculty and other data with respect to the various psychological institutes. In those universities where additional institutes have been added, the figures are likely to be conservative. Not all of these institutes gave sufficient information concerning the new faculty, students, and facilities which had been added. Table 1 also makes it possible for the reader to note the growth which has occurred in psychology over a 10-year period. The last column in the table reflects the percentage of increase, with the 1955-1966 data as a base. It becomes apparent that there has been phenomenal growth when these 21 universities are considered as a unit. Two of the items in Table 1 may need further explanation. The item reflecting the number of students enrolled in psychology refers to students that are currently taking psychology courses. The "number of mechanics employed" will not be a familiar category for American psychologists who are used to constructing their own research equipment. At Humboldt University, two of the eight mechanics employed are in the field of electronics.

In 1941 the Diplom degree was introduced into Germany. Austria does not have a similar degree. This degree might be thought of as an exceedingly difficult Master's degree. It is possible to complete the doctorate in one or two additional semesters. The Diplom degree was introduced to meet the need for applied psychologists. A reform of this degree was introduced in 1955 for uniformity in the West German universities; this standardized eight semesters of study and the *Vorprufung,* or examination, for the degree. At the same time, in East Germany, a reform was introduced which extended the Diplom training to 10 semesters. If the increase in the numbers of students that are receiving this degree can be taken as indication of its success in West Germany, it is evident that this practitioner qualification is meeting a real need.

The number of doctoral students has also increased; although not as dramatically as the surge in Diplom students. The University of Vienna continues to have the largest number of doctoral students in both Austria and Germany, while Göttingen has the largest number in Germany. The growth in the physical space of the institutes appears to be paralleling the growth in faculty and students. Münster is now the university with the most physical space (if number of rooms is a reliable indication), and the universities in Berlin, Göttingen, and Bonn follow closely behind.

F. Summary

This survey of the status of psychology within the universities of Austria and Germany has revealed some new developments. A number of psychological institutes have been added to those already in existence within the West German universities. One function of these institutes is to make available psychological training for students whose primary training is being conducted in other faculties. It may also indicate that a greater emphasis is being placed on the applied aspects of psychology than existed before. It should be of particular interest to note the future direction of research within these universities.

Research trends over the last 10 years have also been examined. It seems that test construction research is waning; statistics, mathematical models, and experimental design research is increasing; developmental psychological research is on the decrease; and the traditional areas of psychological concentration continue at about the same pace.

Over 4000 students are currently enrolled in the German diploma programs —almost a 300 per cent increase within 10 years. This seems to indicate that this practitioner degree is continuing to meet a need. Growth in facilities, doctoral students, and faculty advances at a steady pace. All in all, the status of psychology within the universities of Austria and Germany seems to be secure and full of vitality.

References

1. Adams, J. F. The status of psychology in the universities of Austria and Germany: 1955-1956. *J. Gen. Psychol.,* 1957, **56**, 147-157.

2. ———. Towards better communication of psychological research between countries. *J. Gen. Psychol.,* 1959, **60**, 289-290.
3. ———. The status of psychology in the universities of Austria and Germany: 1960-1961. *J. Gen. Psychol.,* 1962, **67**, 337-347.
4. DAVID, H. P., *et al.* Reciprocal influences in international psychology. *Amer. Psychol.,* 1960, **15**, 313-315.
5. MCCOLLUM, I. N. Psychologists in industry in the United Kingdom and Western Germany. *Amer. Psychol.,* 1960, **15**, 58-64.
6. METZGER, W. The historical background for national trends in psychology: German psychology. *J. History Behav. Sci.,* 1965, **1**, 109-115.
7. RAZRAM, G. Growth, scope, and direction of current Soviet psychology. *Amer. Psychol.,* 1964, **19**, 342-349.
8. WATSON, G. Psychology in Germany and Austria. *Psychol. Bull.,* 1934, **31**, 755-766.
9. WESLEY, F. Masters and pupils among German psychologists. *J. History Behav. Sci.,* 1965, **1**, 252-258.

Department of Psychology
Temple University
Philadelphia, Pennsylvania 19122

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 135-141.

SUBSUMPTIVE LABELING AS A VARIABLE IN CONCEPT ORGANIZATION*

Department of Psychology, Utah State University

DAVID R. STONE

A. INTRODUCTION

There is currently an increased interest in cognitive processes including problem-solving as a form of complex human behavior [see Staats and Staats (2) and Harper *et al.* (5)].

The influence of prior experience, as reflected in verbal mediators, is one related problem area [see Luria and Yudovich (7)]. In other words, acquired language patterns influence the interpretation of new events.

The varied use of verbal categories can influence the solution of a problem. Rees and Israel (10), for example, used anagrams (mixed letters) in a task in which the rearranged letters formed "nature words." Later, they used words which could be formed into either "nature" or "food" words. Results revealed that subjects who had experience with one verbal form continued to use the form in solving future problems, more than a control group lacking such prior experience.

The experimental question is: Do the kinds of labels available to students influence the quality of their solutions to problems?

The aim of the present research is to observe the effect of some verbal organizers that are arranged logically to include others. This process of subsumption is given by Ausubel (1) as a basic research area. When the subjects are confronted with variable stimuli, what factors influence the way in which these subjects make or try to "make sense" out of them?

The Welch study (14) showed that children are less able to learn a higher order concept ("animal" in an "animal-cat-Siamese" sequence) than they are to learn a lower order concept. Welch contended that such higher order abstractions, being farther from the direct stimulus, are harder to learn than lower order abstractions.

This kind of category structure means that a given stimulus may be classified at various levels of generality, using any or all of the terms; and may be classified at more and more abstract levels by progressively using the

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 5, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

more general terms. This is seen in the terms: kingdom, phylum, class, order, family, genus, and species, used to clarify relationships in zoology.

Although often used in instruction, information given by analogy has not been the subject of controlled study. Studies of verbalization in learning have generally followed three major trends: (*a*) the improvement of labeling to distinguish parts, as in Kendler and Vineberg (6), (*b*) the statement of guiding hypotheses, as in Schwartz and Lippmann (11), and (*c*) the organization of techniques, as in Ausubel (1). All three trends are ways of clarifying the relationship of verbal labels to the problem.

The theoretical background for this study is seen in the seemingly opposing views of Bruner's emphasis in self-guided learning (2) and Ausubel's emphasis on didactic learning (1)—in other words, discovery *versus* "being told." One area of compromise appears to be in providing semiguides, hints in sequence, and partial cues, so that the learner may participate inductively in the process, as in Ray (9). While transfer studies [see Orata (8) and Forgus and Schwartz (4)] indicate that direct instruction is most effective in building concepts, there is evidence [see Corman (3)] to suggest that self-directed learning may produce longer retention than direct instruction. Is it possible to obtain cross-benefits by combined procedures? This study attempts to begin the evaluation of one such method—that of analogous cueing. The purpose of the experiment is to investigate how partial cues aid in arranging concepts in an hierarchic or subsumptive fashion.

The basic hypothesis is: students who are given a representative set of subsumptive labels will order random stimuli more successfully than those who have no such system.

B. METHOD

The subjects were 313 upper division university students, male and female. They were divided into two groups, at random, for the experiment.

The task was to identify the relationship between seven nonsense syllables and four geometric figures. The nonsense syllables were DAX, MEF, TOV, VIC, YOP, ZIL, and ZIP. The figures used were an obtuse triangle, a right triangle, a rectangle, and a square.

In order to solve the task the subject had to discover that ZIP applied to all the figures, DAX to the right angle figures, and VIC to the triangles. The remaining syllables were the specific names for each figure. In the experiment, the words and the figures were presented in a preset schedule to be described later.

Learning of this type, while basic to scientific taxonomy, has not been explored.

For the experimental treatment, prior to the paired presentation series half of the subjects were given seven simple terms that illustrated by analogy the general principle of subsumption. These terms were: "animal," which covered all cases; and subcategories "wild" and "tame," with "lion" and "tiger" under "wild" and "cat" and "dog" under "tame." To the degree that the subjects realized the seven terms could be used to categorize the four particular animals, it was hypothesized that they would also be able to see the relationship between the four geometric figures and the seven nonsense terms.

The purpose of the practical application of the procedure was to illustrate, by analogous cueing, that subjects could have an experience of "discovery learning" more valuable than direct telling. The purpose of the research application was to explore how the degree of prompting [see Stolurow (13)] influences learning efficiency and learning attitudes.

The sequence in Figure 1 will serve as an analogous cue-system for the relationship to be found in the experiment.

Figure 2 shows the subsumptive relationships presented to subjects as pairs.

Animal
Wild
Lion
Tiger
Tame
Dog
Cat

FIGURE 1
PRERESPONSE ANALOGY GIVEN TO EXPERIMENTAL GROUP

ZIP—The general set or class name for all figures.
VIC—The subset name for triangles.
TOV—The obtuse triangle.
MEF—The right triangle.
DAX—For right-angle figures.
YOP—The square.
ZIL—The rectangle.

FIGURE 2
THE RELATIONSHIP OF NONSENSE SYLLABLES USED AS NAMES FOR THE GEOMETRIC FIGURES

The subjects' task was to learn the nonsense-syllable labels for the general set, the subsets, and the four specific geometric figures.

One-half of the subject group was kept out of the room during the experimental instructions, thus serving as a control.

The experimenter then gave the following instructions to both groups:

> In a moment we are going to present a group of four geometric figures cut out of black cardboard. The object is for you to figure out the naming or labeling relationship between the seven nonsense words and the four figures. All are used. With each presentation we will give one "bit" of information. For example, the first bit (holding up the square) is: This is a DAX. We will give you 12 bits, then pause two minutes while you write down on the record sheet any conclusions you have made. Then we will give you 12 more bits, and then allow five minutes for you to write the observations you can about the relationship of the words to the figures. Write the words on the upper left side of the data sheet for reference. Show the figures across the top of the page, also for reference. Keep notes on each bit of information as it is given.

As shown in Table 1, the first presentation was the square given with the statement, "This is a DAX." The others were given in order as listed. Nega-

TABLE 1
SEQUENCE OF NONSENSE SYLLABLES USED AS NAMES FOR GEOMETRIC FIGURES

Sequence	Figure	Specific	Subset	General set
1	Square		+DAX	
2	Rectangle			+ZIP
3	Obtuse triangle		+VIC	
4	Right triangle			+ZIP
5	Square	−TOV		
6	Rectangle		−VIC	
7	Obtuse triangle	−ZIL		
8	Right triangle		−DAX	
9	Square	+YOP		
10	Rectangle	−MEF		
11	Obtuse triangle	+TOV		
12	Right triangle	−YOP		
13	Square	−MEF		
14	Rectangle	+ZIL		
15	Obtuse triangle	−YOP		
16	Right triangle	+MEF		
17	Square		−VIC	
18	Rectangle	−TOV		
19	Obtuse triangle		−DAX	
20	Right triangle	−ZIL		
21	Square			+ZIP
22	Rectangle		+DAX	
23	Obtuse triangle			+ZIP
24	Right triangle		+VIC	

tive signs in Table 1 indicate figures given with the statement, "This is *not* a ———."

The students kept their own records as the pairs were presented. These records were later rated as to coverage and organization.

Table 2 shows what each subject could know at the end of the experiment if his system was complete. Note that four bits of information were withheld (0), so that some inferential thinking was required at the specific object level. For example, by comparing the double reference for the square and for the rectangle, the individual could conclude that both were DAX, and (by bit 23) infer that the square was YOP and the rectangle ZIL. These incomplete bits were used to provide a somewhat broader response range than a completely cued program would have done.

C. RESULTS

Scoring was based on the number of correct conclusions as to the placing of the seven labels in the hierarchy. With the use of an IBM computer, a three-way independent chi-square analysis was carried out, based on the data in Table 3.

Significant differences were found (one per cent level) for cued *vs.* non-

TABLE 2
A SUMMARY OF INFORMATION GIVEN FOR EACH GEOMETRIC FIGURE

Nonsense syllables	Square	Rectangle	Obtuse triangle	Right triangle
DAX	+	+	—	—
MEF	—	—	0	+
TOV	—	—	+	0
VIC	—	—	+	+
YOP	+	0	—	—
ZIL	0	+	—	—
ZIP	+	+	+	+

TABLE 3
NUMBER OF ELEMENTS CORRECT IN CUED AND NONCUED GROUPS

Cues given	Note quality	Elements correct								Total
		0	1	2	3	4	5	6	7	
Cued	S+	5	12	5	25	3	3	14	51	118
	S—	4	3	5	8	3	7	1	8	39
Noncued	S+	7	9	6	9	6	18	8	20	83
	S—	10	11	7	11	8	8	6	12	73

cued, and for plus *vs.* negative data systems in general. Plus systems were those in which the data could be readily used because the student had recorded it in a good style or form.

It was clear that subjects who had received preresponse cueing by analogy had produced more useful notes, and had achieved more and better solutions than subjects who did not receive preresponse cueing.

D. Discussion

The order of presentation was arranged so that the final use of the third-level label, ZIP, was left until pair number 22. Final verification of the second-level labels was also delayed until the last three pairs.

Obviously, variations in timing and sequence of label-levels should be the subject of future experimentation. The general question is raised as to how the early presentation of second- and third-level labels would facilitate learning as compared to other possible sequences.

E. Summary and Conclusions

In this experiment, with 313 university students as subjects, a series of nonsense syllables (DAX, MEF, TOV, VIC, YOP, ZIP, and ZIL) were to be associated with four geometric figures. The task was so arranged that ZIP applied to all figures, DAX and VIC to subsets of two figures each, and the remainder were individual labels. One group received preresponse cueing by means of an analogy which involved hierarchic concepts in the same general form: i.e., animal, wild, tame, and individual names.

A significant number (one per cent level) of the preresponse cued groups achieved a superior solution. They were also able to set up better systems for handling the data than were members of the control group.

References

1. Ausubel, D. P. The Psychology of Meaningful Verbal Learning. New York: Grune & Stratton, 1963.
2. Bruner, J. S. The Process of Education. Cambridge, Mass.: Harvard Univ. Press, 1960.
3. Corman, D. The effect of varying amounts and kinds of information as guidance in problem solving. *Psychol. Monog.*, 1957, **71**, Whole No. 431.
4. Forgus, R. H., & Schwartz, R. J. Efficient retention and transfer as affected by learning methods. *J. of Psychol.*, 1957, **43**, 135-139.
5. Harper, J. C., *et al.* The Cognitive Processes: Readings. Englewood Cliffs, New Jersey: Prentice-Hall, 1964.
6. Kendler, H. H., & Vineberg, R. The acquisition of compound concepts as a function of previous training. *J. Exper. Psychol.*, 1954, **48**, 252-258.

7. LURIA, A. R., & YUDOVICH, F. Speech and the Development of Mental Processes in the Child. London: Stables, 1959.

8. ORATA, P. T. Recent research studies on transfer of training with implication for the curriculum, guidance, and personnel work. *J. Educ. Res.*, 1941, **35**, 81-101.

9. RAY, W. S. Verbal compared with manipulative solution of an apparatus problem. *Amer. J. Psychol.*, 1957, **70**, 289-290.

10. REES, H., & ISRAEL, H. An investigation of the establishment and operation of mental sets. *Psychol. Monog.*, 1935, **46**, Whole No. 210.

11. SCHWARTZ, F., & LIPPMANN, F. Cognitive and associative structures in recall. *Psychol. Rep.*, 1962, **11**, 91-101.

12. STAATS, H. W., & STAATS, C. K. Complex Human Behavior. New York: Holt, Rinehart & Winston, 1964.

13. STOLUROW, L. M. Psychological and Educational Factors in Transfer of Training, Nos. 8 and 9, University Microfilms, Ann Arbor, Michigan, 1965.

14. WELCH, R. L. A preliminary investigation of some aspects of the hierarchal development of concepts. *J. Genet. Psychol.*, 1940, **56**, 359-378.

Department of Psychology
Utah State University
Logan, Utah 84321

Published as a separate and in *The Journal of Psychology,* 1966, **63**, 143-156.

DIFFERENCES BETWEEN ACCEPTED AND REJECTED PATIENTS IN A TREATMENT PROJECT OF COLLEGE DROPOUTS*[1]

The William Alanson White Institute

MARTIN KOHN AND EDGAR A. LEVENSON

A. INTRODUCTION

For the past two years the William Alanson White Institute of Psychiatry, Psychoanalysis and Psychology has conducted a demonstration treatment program geared to the intellectually capable college dropout. During this period a total of 90 patients have been completely screened; of this total, 60 per cent were accepted for treatment and 40 per cent were rejected. This paper presents a comparison of these two groups.

The major aim of the present paper is to examine (*a*) the extent to which the selection of patients was in line with stated intake criteria of the Institute, and (*b*) the extent to which psychological as opposed to demographic variables differentiated between accepted and rejected patients. Previous research (2, 3, 5, 7, 8, 9, 12) has focused on relatively gross indices such as sex, age, education, etc., while psychological variables have been relatively neglected. Specifically, the authors were interested in discovering to what extent such factors as a patient's optimism, rebelliousness, and striving for achievement implicitly governed the Institute's selection.

The details of the project have been described extensively elsewhere (6). Briefly stated, its purpose was to establish a community clinic for college dropouts, which would function both as a treatment center and as a focus for research and continuous data collection. The students were seen only after their separation from the campus. It is to be emphasized that the staff saw only dropouts suspected of having a significant emotional component in their

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 11, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This project was supported by Grant MH-867 from the National Institute of Mental Health to the William Alanson White Institute of Psychiatry, Psychoanalysis and Psychology. The authors are indebted to Mrs. Ruth Tendler, Mrs. Marjorie Newstrand, and to the other participants in the project for their sustained efforts in collecting the data, and to Mrs. Judith Kraemer for her valuable work in the analysis of these data.

inability to continue in school, and who were regarded by their college as potentially competent students. Individual psychotherapy, once or twice weekly, and group therapy as an adjunct were offered for an optimum period of up to one year, although patients were continued into the second year if warranted.

Screening of the applicants consisted of (*a*) filling in an application form, (*b*) an initial interview with the chief psychiatrist, (*c*) an interview with the social worker, (*d*) a psychological battery of tests (including the Rorschach), and (*e*) filling in a series of specifically selected research forms. With the exception of the research forms, this material was reviewed in a screening conference which included the psychiatrist, the psychologist, the social worker, and any other staff members interested in attending. Acceptance for therapy was assessed according to three main criteria: first, the extent to which the staff estimated that the dropping-out was causally connected with the college experience as an emotional crisis; second, an assessment of the student's intelligence, originality, and genuine interest in education; and third, an assessment of the likelihood of his responding to therapy. Severity of pathology was, if possible, not weighed heavily; nor were patients selected simply because workers thought that they would be likely to return to school.

The first part of this paper presents differences on data which were utilized in the screening conference: namely, the data from the psychiatric interview, the social work interviews and the Rorschach test. The second part presents data not utilized in arriving at the decision of accepting *vs.* rejecting the patient: namely, the results from Strodtbeck's V-Scale (10), a test of achievement striving, and a specially designed sentence completion test.

B. THE SAMPLE

Colleges likely to have students from the New York metropolitan area were contacted and informed of the project. They were asked to refer to the Institute dropouts who seemed potentially capable students and in whom some degree of emotional disturbance was suspected.

It may be useful to delineate briefly the salient characteristics of the total patient group referred to the project:

1. The median age was 19, ranging from 17 to 26.

2. The median *IQ* was 121, ranging from 102 to 141 (75 per cent of the patients had *IQ*s above 115).

3. Thirty-six per cent of the patients dropped out during their first year of college, 31, 25, and eight per cent during their second, third, and fourth year respectively.

4. Ninety per cent of the group dropped out of college more or less

precipitously; only 10 per cent completed the year in relatively good standing.

5. The patients came from a wide range of socioeconomic backgrounds: 40 per cent of the fathers had a high school education or less, 37 per cent had some college or had graduated from college, and 33 per cent had done some graduate work or had completed a course of graduate study.

Items 2 and 4 above suggest that the Institute dealt with a particular section of the dropout group: i.e., patients who were relatively bright and whose dropping out of college was precipitated by considerable emotional stress.

The 90 patients who completed screening, and who form the subject of this paper, divided themselves as follows: Accepted Group: 54 patients (Males, 34; Females, 20); Rejected Group: 36 patients (Males, 17; Females, 19).

Since the two groups contain unequal proportions of males and females, tests for significance of male-female differences were made; only those results that are unlikely to be due to a male-female differential are presented in the present paper.[2] Specific evidence in this respect will be presented in a later section of the paper. All tests of significance were carried out on a two tailed basis. Tests of significance were carried out by chi-square analysis (4).

C. The Intake Interviews

As a first step in the screening procedure, all applicants to the project were seen in initial interviews by the project psychiatrist and the project social worker. To utilize the data from the interviews a schedule[3] was prepared; the data fell into the six major clusters presented in Table 1. Analysis of the data focused on the extent to which accepted patients differed from the rejected patients in (*a*) socioeconomic background, (*b*) their initial relationship to the clinic, (*c*) intellectual endowment and achievement, (*d*) gross pathology of family members, (*e*) adequacy of peer relations, and (*f*) destructiveness of their parents' attitudes towards their life plans.

As seen in Table 1, very few of the differences between the accepted and the rejected patients on the identifying, demographic, and background variables

---

[2] An alternative would have been to remove cases randomly in order to achieve equal proportions of males and females; however, this would have reduced the size of the sample considerably.

[3] The information culled from the interviews and recorded on the schedules was subjected to tests of interscorer agreement. On most of the items the agreement between first and second scoring was 80 per cent or more. (See Kohn, M., and Levenson, E. A., "Some Characteristics of a Group of Bright, Emotionally Disturbed College Dropouts," *Coll. Health,* 1964, **12** (No. 4), 382-391. This paper also presents detailed definitions of the categories of analysis.)

TABLE 1
INITIAL INTERVIEW DATA

Variables assessed in initial interview	Accepted patients	Rejected patients	Statistical significance
I. Identifying and demographic variables			
1. Median age (in years)	19	20	N.S.
2. From New York City colleges	61%	72%	N.S.
3. Mode of separation: dropped by college	36%	33%	N.S.
4. Lived with parents during college	47%	65%	.05
5. Lived with parents after dropout	68%	65%	N.S.
6. Education—Fathers:			
Some high school or completed	32%	49%	N.S.
Some college or completed	67%	51%	N.S.
7. Occupation—Fathers:			
Professional or semiprofessional	52%	46%	N.S.
Small business	15%	11%	N.S.
Clerical	8%	9%	N.S.
Skilled, semiskilled or unskilled	23%	26%	N.S.
Unusual	2%	8%	N.S.
8. Education—Mothers:			
Some high school or completed	43%	54%	N.S.
College	52%	47%	N.S.
9. Occupation—Mothers:			
Lesser professional or small business	19%	36%	N.S.
Clerical	33%	27%	N.S.
Service, skilled, or semiskilled	13%	0%	N.S.
Unusual	9%	5%	N.S.
Housewife only	26%	32%	N.S.
II. Contact with clinic			
1. Median time lag (in months) between dropout and application	2.2	5.7	.05
III. *IQ* and academic achievement			
1. Median *IQ*	123	119	N.S.
2. College entrance examination:			
Higher than 600	58%	28%	.05
Wide scatter	18%	40%	N.S.
3. College academic record year of dropout:			
Superior record	10%	16%	N.S.
Failing record	27%	34%	N.S.
IV. Aspects of family pathology			
1. From broken homes (divorce, separation, or death of a parent)	32%	31%	N.S.
2. Broken home where broken by death of a parent	44%	18%	N.S.
3. Gross disturbance of family members:			
Father	23%	37%	N.S.
Mother	10%	31%	.05
Siblings	5%	27%	N.S.

TABLE 1 (*continued*)

Variables assessed in initial interview	Accepted patients	Rejected patients	Statistical significance
V. Peer relations in adolescence			
1. Same sex social competence:			
Low	24%	41%	N.S.
Fair	51%	38%	N.S.
High	24%	21%	N.S.
2. Opposite sex social competence:			
Low	33%	54%	N.S.
Fair	57%	33%	N.S.
High	10%	12%	N.S.
VI. Parents' involvement in patient's life plans			
1. Mother's ambition for patient:			
Pushing	70%	52%	N.S.
Lack of interest	11%	22%	N.S.
Realistic ambition	11%	13%	N.S.
Unrealistic ambition	8%	13%	N.S.
2. Father's ambition for patient:			
Pushing	55%	36%	.05
Lack of interest	21%	50%	.05
Realistic ambition	18%	5%	.05
Unrealistic ambition	6%	9%	.05

are significant. Differences between the two groups reaching significance were:

1. More of the rejected patients than of the accepted patients lived with their parents while going to college.
2. The accepted patients were higher than the rejected patients on the College Board Entrance Examination.
3. Fewer of the accepted patients than of the rejected patients had mothers with gross emotional disturbance.
4. The fathers of the accepted patients showed a pattern of destructiveness different from the pattern shown by the fathers of the rejected patients; more of the latter were "pushing" whereas more of the former showed lack of interest.
5. The rejected patients were significantly slower in applying to the clinic after dropping out than the accepted patients.

D. RORSCHACH FINDINGS

In making the decision to accept or reject a patient, the Institute relied heavily on the Rorschach test. A systematic analysis of the Rorschach responses was carried out. Practical limitations precluded carrying out an analysis on the interpretive level; therefore, differences between the accepted and rejected groups on the uninterpreted Rorschach variables are presented (see Table 2).

Three of the major Rorschach variables reach the five per cent level of significance:

1. The accepted patients gave significantly more whole responses (W) than the rejected patients.

2. The accepted patients gave significantly more animal movement responses than the rejected patients.

TABLE 2
RORSCHACH VARIABLES

Rorschach response categories	Accepted patients	Rejected patients	Statistical significance
More than 31 responses	51%	49%	N.S.
Whole responses (W's), more than 8	60%	37%	.05
Large detail (d%), above median, 54%	42%	64%	.05
Human movement (M), 3 or more	55%	46%	N.S.
Animal movement (FM), none	10%	23%	.05
Inanimate movement (m), none	26%	46%	N.S.
Sum of color responses, 3.5 or more	55%	40%	N.S.
F plus %, above median of 74, 5%	57%	43%	N.S.
Human responses, 5 or more	23%	17%	N.S.
Animal responses (A%), above median (—40%)	49%	47%	N.S.
Content in more than 11 categories	58%	41%	N.S.
Space responses, none	32%	52%	N.S.
Original responses, 2 or more	55%	37%	N.S.
Popular responses, 6 or more	64%	66%	N.S.
No *initial* turns	62%	46%	N.S.
Total turns other than initial turns	50%	49%	N.S.

3. The rejected patients gave significantly more large detail responses than the accepted patients. Since the overall Rorschach productivity of the two groups was the same, the larger amount of whole responses and the smaller number of large detail responses of the accepted group than of the rejected group are two sides of the same coin.

E. STRODTBECK'S V-SCALE

This scale consists of 14 true-false statements, which as a whole correlate highly with achievement-striving. In a factor analysis of the scale, Strodtbeck (10) isolated two factors: (*a*) mastery and (*b*) independence from family. The scale was administered because of the relevance of achievement-striving to dropout dynamics, and possibly to eagerness to utilize therapy.

As originally administered by Strodtbeck, each item was marked as true or false; to achieve finer discriminations the patients were asked to indicate their degree of agreement on a six-point scale from minus three indicating strong disagreement to plus three indicating strong agreement.

The results are presented in Table 3. When the median value of the entire group (+19) was used as a cutoff point, 63 per cent of the accepted group as compared to 35 per cent of the rejected group fell above it. The accepted group, in other words, showed significantly higher achievement-striving than the rejected group.

TABLE 3
V-SCALE

Scores derived from V-Scale	Accepted patients	Rejected patients	Statistical significance
Total score (14 scales), score of 19 or more	63%	35%	.05
Independence from family (3 scales), score of 8 or more	90%	69%	.03
Mastery (2 scales), score of 4 or more	60%	62%	N.S.

A further analysis was carried out on the items that had high factor loading on "independence from family" and "mastery." On the "independence from family" items, the accepted group was significantly higher than the rejected group. There was, however, only a slight, nonsignificant difference between the two groups on the "mastery" items.

F. SENTENCE COMPLETION

A sentence completion test was developed by the Institute; it was composed of four groups of stems designed to elicit the patient's attitude toward: (*a*) the future, (*b*) the past, (*c*) dealing with obstacles, and (*d*) dealing with criticism.[4]

Each response was classified into its relevant category (see Appendix A) or as a direct evasion. We were interested in evasive ways of responding, since previous research has shown that this measure differentiates "rebels" from "conformists" and "independents" (11). Examples of evasive responses are "tomorrow . . . is tomorrow," or "As I grow up . . . I increase in age."

Responses were scored on the basis of manifest content; no attempt at interpretation was made. Also, so far as possible each answer was scored completely independently without taking into account other completions. Two scorers working independently on the records from 10 patients showed 85 per cent agreement.

---

[4] The incomplete sentence stems as well as the categories of analysis can be found in Appendix A.

In the first and the second parts of Table 4, the data on the patient's orientation to the future and the past are presented.

Two of the differences reached statistical significance:

1. With regard to the future, the rejected patient showed a significantly stronger negative attitude than the accepted patients ($p = .05$).

TABLE 4
SENTENCE COMPLETION TEST

Response categories	Percentage of patients giving one or more responses Accepted patients	Rejected patients	Statistical significance
1. Attitude toward future			
Positive-mild	42	36	N.S.
Positive-strong	52	39	N.S.
Negative-mild	12	21	N.S.
Negative-strong	65	86	.05
2. Attitude toward past			
Positive	88	75	N.S.
Negative	88	100	N.S.
More negative than positive	38	64	.02
3. Dealing with obstacles			
Appropriate response	42	43	N.S.
Refuse to admit defeat	50	71	N.S.
Seeks help	12	11	N.S.
Quits	67	54	N.S.
Self-hostile	56	75	N.S.
Hostile to others	23	25	N.S.
4. Dealing with criticism			
Appropriate response	33	46	N.S.
Hostility to others	23	18	N.S.
Hostility directed to self	40	61	N.S.
Escape, acting helpless or passive response	60	43	N.S.
Hiding or denying feelings	40	50	N.S.
5. Evasive response			
Re: orientation to future	46	39	N.S.
Re: orientation to past	63	39	.05
Re: dealing with obstacles	2	0	N.S.
Re: dealing with criticism	2	0	N.S.

2. With regard to the past, the rejected patients showed a significantly stronger negative attitude than the accepted patients ($p = .02$).

Parts 3 and 4 of Table 4 present the results on "Dealing with obstacles" and "Dealing with criticism." None of the differences reached statistical significance, although on both "Dealing with criticism" and "Dealing with obstacles" the rejected patients had a greater tendency to respond with self-hostile completions than the accepted patients.

The last part of the table presents the results on "Evasive responses." The table shows one significant difference. In their completion of the "Orientation to the past" stems, the accepted patients gave a significantly larger number of evasive responses than the rejected patients.

G. DISCUSSION

Since the accepted and the rejected patients contain different proportions of males and females, it is important to document that none of the differences between the two groups can be accounted for by a male-female differential on these variables. All variables found to be significant at the five per cent level or better were examined for male-female differences. The results, presented in Table 5, show that none of the male-female differences was significant.

TABLE 5
DIFFERENCES BETWEEN MALE AND FEMALE PATIENTS

Data analysis categories	Males $N = 51$	Females $N = 39$	Statistical significance
Initial interview data			
Lived with parents during college	54%	54%	N.S.
Time lag between dropout and application	3.4 mos.	3.6 mos.	N.S.
Some disturbance of family members—mother	23%	14%	N.S.
Father's ambition for patient			
Pushing	55%	37%	N.S.
Lack of interest	36%	30%	N.S.
Realistic ambition	6%	21%	N.S.
Unrealistic ambition	3%	12%	N.S.
Rorschach variables			
Whole responses (W's), more than 8	41%	33%	N.S.
Large detail (d%), above median 54%	55%	44%	N.S.
Animal movement (FM), none	16%	13%	N.S.
V-Scale			
Total score (14 scales), score of 19 or more	53%	43%	N.S.
Independence from family (3 scales), score of 8 or more	55%	45%	N.S.
Sentence completion			
Negative-strong	70%	75%	N.S.
More negative than positive	45%	50%	N.S.
Evasive responses	42%	42%	N.S.

In each of the four sets of data some differences between the two groups have been found. To provide an overview, differences significant or approaching significance will be briefly summarized below. On the left side of Table 6 are the findings; on the right side, their implications for personality functioning.

The following is suggested when the findings of this study are related to the stated intake criteria of the project: contrary to the project's intention not to make degree of disturbance a major consideration, the rejected patients seem to be a more poorly functioning group than the accepted patients. The accepted patients, as compared to the rejected patients, showed (*a*) greater initiative in resolving their difficulties, (*b*) more achievement-striving, (*c*)

TABLE 6
SUMMARY OF DIFFERENCES BETWEEN ACCEPTED AND REJECTED PATIENTS

Findings	Implications for personality functioning
Intake interview:	
Time lag between dropout and application shorter for the accepted group than for the rejected group ($p = .05$).	Accepted group showed greater initiative in resolving difficulties than rejected group.
Accepted patients had higher college board examination scores than rejected patients ($p = .05$).	Accepted patients were better prepared for college or better able to use their resources than rejected patients.
Accepted group had fathers who were somewhat more pushing and had more realistic ambitions for them, fewer fathers showing lack of interest and unrealistic ambitions than rejected group ($p = .05$).	Fathers of accepted patients showed a different pattern of deviant father-child relationship, showed a somewhat greater involvement and possibly more concern with the child than fathers of rejected patients—even though such concern may be highly deviant.
Rorschach:	
Rorschach scores of the accepted patients showed more W, FM than scores of rejected patients ($p = .05$, $.05$).	Accepted patients gave more expansive records, showed more ambition and drive than rejected patients.
Strodtbeck's V-Scale:	
Accepted patients had higher total "V" score than rejected patients ($p = .05$).	Accepted patients were more achievement oriented than rejected patients.
Accepted patients were higher on "Independence from family than rejected patients ($p = .03$).	Accepted patients were more willing to sever ties from parents to achieve independence than were rejected patients.
Sentence completion:	
Accepted patients gave more positive responses to stems dealing with past and future than rejected patients ($p = .05$ to $.02$).	Accepted patients were generally more optimistic than rejected patients.
Accepted patients were more likely to make more direct evasions in response to incomplete stems than rejected patients ($p = .05$).	Accepted patients were a more rebellious group than rejected patients.

greater conscious willingness to separate themselves from their parents and (*d*) a more positive conception of both the past and the future.

At the same time and largely for the same reasons, there is some slight evidence that the accepted patients may be more in line with our second criteria than the rejected patients: they seem to be better educational bets than the rejected group. As judged by their college entrance scores, they are better prepared educationally than the rejected patients. Whether they would translate this potential into achievement is, however, another question. There is some suggestion in the data that they may; as indicated previously their achievement-striving is higher and they seem to be more willing to act in their own behalf than the rejected group.

According to its third intake criteria, the project wanted to select the patients who would show the greatest likelihood of responding to therapy. The relatively better personality functioning of the accepted patients over the rejected group suggests that they may be able to benefit from treatment more rapidly than the rejected patients. However, in selecting them the Institute at the same time "unwittingly" chose a group that is more rebellious than the rejected group. (See Sentence Completion Test.) This personality trait may be expected to increase the difficulties in the therapeutic transactions; it may in fact account for the complicated transference and counter-transference relationship that did in fact develop in the work with these patients (1).

Data from the two instruments not utilized during screening—the V-Scale and the sentence completion test—suggest that unwittingly the project chose patients who were higher on achievement striving, who were more optimistic, more rebellious and tended to be less self-hostile than the rejected group. These differences were by and large more decisive than the differences on the demographic variables. The results suggest that the gross demographic variables, particularly education and occupation, that have frequently been found to differentiate between accepted and rejected patients, may simply be crude correlates of subtle psychological variables which screeners feel to be of importance in selecting patients.

H. SUMMARY

The study examined differences between accepted and rejected patients in a clinical project on college dropouts. The patients were compared on data from (*a*) psychiatric and social work interviews, (*b*) the Rorschach Test, (*c*) Strodtbeck's V-Scale, and (*d*) a sentence completion test.

The results showed that in comparison with the rejected patients, the ac-

cepted patients were more achievement oriented, more willing to separate from their families, more positive both with respect to the past and the future, and more rebellious. These psychological variables tended to differentiate the two groups more sharply than the background variables such as education and occupation. The findings suggest that gross variables, such as socioeconomic level, that have frequently been found to differentiate between accepted and rejected patients may in fact be only crude indicators of subtle psychological variables which screeners take into account.

APPENDIX A

A. Sentence stems

1. Orientation to the future:
    - As you grow older . . .
    - Next month . . .
    - When he thinks of tomorrow . . .
2. Orientation to the past:
    - It used to be that . . .
    - To his parents he was . . .
    - In the old days . . .
    - When he was a child . . .
    - In high school . . .
    - At home . . .
3. Dealing with obstacles:
    - If the job is too hard . . .
    - When things went against him . . .
    - When he can't do it . . .
    - It looked impossible so he . . .
4. Dealing with criticism:
    - In the face of criticism . . .
    - When the prof bawled him out . . .
    - When they laughed at him . . .

B. Categories of analysis

1. Orientation to future:
    - a. Positive: general statements indicating hopefulness and optimism
    - b. Positive: definitive positive statements: e.g., anticipation of completion of step towards goals, expansive statements, or

statements implying gains of self-development or in relations to world

- c. Negative: mild, moderate, or extremely negative statements; statements indicating that things will be bad in the immediate or long-range future

2. Orientation to past:
    - a. Positive: things were good, happy, successful in past statements in which past is idealized, personal or impersonal expression of nostalgia
    - b. Negative: mild, moderate, or extremely negative statements; statements indicating that things were bad; expressions of feelings of isolation and loneliness
    - c. Complex: things were both good and bad
3. Dealing with obstacles:
    - a. Objective attitude; appropriate or realistic effort in dealing with obstacle
    - b. Refusal to admit defeat or expression of continual unrealistic effort
    - c. Seeking help or support
    - d. Quitting or expressions of impossibility of successful coping blithely accepted
    - e. Hostility or anger directed against the self: mild, moderate, or extremely self-punitive statements
    - f. Hostility directed against outside world: mild, moderate, or extreme expression of hostility against others
4. Dealing with criticism:
    - a. Objective attitude and appropriate response
    - b. Hostility directed against self; aiding with the aggressor against the self
    - c. Hostility directed against others (hostility, anger resentment, fighting back)
    - d. Escape; acting helpless or passive in face of criticism
    - e. Hiding and denying feelings

REFERENCES

1. AKERET, R. U., & STOCKHAMER, N. The value of peer group supervision in understanding countertransference reactions to college dropouts. *Amer. J. Psychother.* (in press).
2. BRILL, N. Q., & STORROW, H. A. Social class and psychiatric treatment. *Arch. Gen. Psychiat.,* 1960, **3**, 340-344.

3. Cole, N. J., Branch, C. H., & Allison, R. B. Some relationships between social class and practice of dynamic psychotherapy. *Amer. J. Psychiat.*, 1962, **118**, 1004-1012.
4. Edwards, A. L. Statistical Analysis. New York: Rinehart, 1947.
5. Hollingshead, A. B., & Redlich, F. R. Social Class in Mental Illness: A Community Study. New York: Wiley, 1958.
6. Levenson, E. A., & Kohn, M. A demonstration clinic for college dropouts. *Coll. Health*, 1964, **12** (No. 4), 382-391.
7. Myers, J. K., & Auld, F. Some variables related to outcome of psychotherapy. *J. Clin. Psychol.*, 1955, **11**, 51-54.
8. Rosenthal, D., & Frank, J. D. Fate of psychiatric clinic outpatients assigned to psychotherapy. *J. Nerv. & Ment. Dis.*, 1958, **127**, 330-343.
9. Schaffer, L., & Myers, J. K. Psychotherapy and social stratification: Empirical study of practice in psychiatric outpatient clinic. *Psychiatry*, 1954, **17**, 83-93.
10. Strodtbeck, F. L. Family interaction, value and achievement. In McClelland, D., *et al. Talent and Society*. New York: Van Nostrand, 1958.
11. Weitman, M. Forms of failure to respond and varieties of authoritarianism. *J. Personal.*, 1964, **32** (No. 1), 109-118.
12. Winder, A. E., & Hersko, M. Effect of social class on length and type of psychotherapy in Veterans Administration Mental Hygiene Clinic. *J. Clin. Psychol.*, 1955, **11**, 77-79.

Research Department
The William Alanson White Institute of Psychiatry,
Psychoanalysis and Psychology
20 West 74th Street
New York, New York 10023

$30.00 per annum
$15.00 per volume
$7.50 per half volume

BIMONTHLY
Three volumes per year
Immediate publication

July, 1966
Volume 63, Second Half

(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a liberal contract, and may not be resold for a period of five years from the date of publication.

JULY, 1966

(Manuscripts are printed in the order of final acceptance)

Rejection of avant-garde art and the need for simple order 159
BY BERNARD PYRON

Runway performance as a function of hunger level measured as food acceptance 179
BY WARREN H. TEICHNER

Assessment of phantom limb sensation among patients with lower extremity amputation 197
BY SAMUEL HIRSCHENFANG AND JOSEPH G. BENTON

Effects of stimulus conditions and response contingencies on the development and maintenance of a perceptual discrimination 201
BY LAWRENCE SIMKINS

Sex differences in the differential recall of taboo and neutral words 219
BY GEORGE S. GROSSER AND ANTHONY A. WALSH

(OVER)

Copyright © 1966 by The Journal Press
Provincetown, Massachusetts, U. S. A.
All rights reserved
Published at 35 New St., Worcester, Mass.
Second-class postage paid at Worcester, Mass.
Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts

Children's behavior problems and mother's age 229
By F. Beth Stone and Vinton N. Rowley

The relationship of confidence to competence and imitation behavior . . . 235
By Herbert Horowitz

Can interview interaction measures be taken from tape recordings? . . . 249
By Arthur N. Wiens, Donald A. Molde, David C. Holman, and Joseph D. Matarazzo

On nonverbal interaction 261
By Joseph Luft

Differential effects of coffee on speed and power tests 269
By G. A. Lienert and H. P. Huber

A comparison between certain retinal and cortical activities underlying vision 275
By S. Howard Bartley

Neuropsychiatric patients, accident proneness, and interest patterns . . . 287
By Joseph T. Kunce and Blayne Brewer

A method for studying olfaction in the unrestrained rat 291
By Warren H. Teichner

Perceived discrepancy in attitude intensity between the actors A and B in *ABX* situations and its effect upon affect 299
By Nehemiah Jordan

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 159-178.

REJECTION OF AVANT-GARDE ART AND THE NEED FOR SIMPLE ORDER*

Instructional Research Laboratory, University of Wisconsin

BERNARD PYRON

A. INTRODUCTION

This study was concerned with personality differences in the acceptance of popular, classical, and avant-garde literature, painting, and music. Most subjects (*S*s) were expected to prefer popular and classical to avant-garde art. The major hypothesis investigated was that *S*s who have a high need for simple order would reject avant-garde art significantly more than those who have a need for complexity.

The human organism seems to need regularity, pattern, and predictability in its stimulus input; but it also needs surprise, novelty, and the unexpected. Some *S*s have a greater need for experiencing novelty, complexity, and ambiguity than for experiencing familiar and easily interpreted sensory input; others, as Budner (4) suggests, have a need to interpret stimuli unequivocally, and tend to reject stimuli that cannot be clearly interpreted within their own relatively simple perceptual organization mechanisms. *S*s who have a greater need for simple order than for complexity would tend to reject unfamiliar or ambiguous art because interpretation would require great effort on their part or because they cannot tolerate uncertainty in the meaning of a work of art.

In this study, the need for simple *vs.* complex order is used as an overriding concept to integrate three subconstructs. The first subconstruct is simplicity-complexity of perceptual organization. A number of studies have dealt with simplicity-complexity of perceptual mechanisms [see Bieri (2), Leventhal (16), Lundy and Berkowitz (17), Mayo and Crockett (18), Scott (30), and Zajonc (35)]. *S*s use at least two kinds of processes to organize stimuli: (*a*) classificatory or nominal organization and (*b*) ordinal or greater-than-or-less-than organization [see Pyron (28)]. Most of the studies of simplicity-complexity of perception have been concerned with ordinal organization. The term "ordering" is used to describe the ordinal organization process (28), since a given stimulus is ordered in relation to other stimuli [see De Soto

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 11, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright, 1966, by The Journal Press.

(5), Helson (9), and Volkmann (34)]. The term "scale" is used to describe the ordinal organizing mechanism rather than (*a*) construct [see Kelly (11)], (*b*) category [see Bruner, Goodnow, and Austin (3)], or (*c*) concept [see Harvey, Hunt, and Schroeder (8)], since stimuli tend to be ordered in terms of the amount of some attribute. The phrase "number of categories" is used to describe the number of discriminable points along a given scale: i.e., a scale may have two or more categories. Simplicity of perceptual organization is maintained by the use of a very small number of different ordering scales and by limitation of the number of categories along the scales.

The second subconstruct is tolerance-intolerance of ambiguity [see Frenkel-Brunswik (7) and Budner (3)]. Budner defines intolerance of ambiguity as "the tendency to perceive (i.e., interpret) ambiguous situations as sources of threat; tolerance of ambiguity is the tendency to perceive ambiguous situations as desirable" (30). Ambiguous stimuli or situations are those characterized by unfamiliarity, novelty, complexity, or insolubility.

The third construct is rigidity. This study was concerned with rigidity as an attitudinal characteristic rather than with rigidity as behavioral perseveration [see Kendig (13), Kounin (15), and Pally (23)]. Attitudinal rigidity (24, 27) is the desire to maintain one's self-concept, beliefs, attitudes, perceptual organization system, and environment in an unchanging condition. Novel or complex stimuli which might threaten or change one's attitudes or perceptual mechanisms are rejected by rigid persons.

Theoretically, one would expect measures of simplicity-complexity of perception, tolerance-intolerance of ambiguity, and attitudinal rigidity to intercorrelate (27, 28). But two factor-analytic studies [see Pyron (28) and Vannoy (33)] show that simplicity-complexity of perceptual ordering is a separate factor. In Vannoy's study, intolerance of ambiguity was part of a separate factor which included independence of judgment and authoritarianism. In a similar study (28), attitudinal rigidity (Acceptance of Change Scale and the Dogmatism Scale) was part of a separate factor which included self-reliance, manifest anxiety, rejection of people, and authoritarianism. Budner (4) suggests that the acceptance of the expectations and norms of others in the culture "reduces the extent of perceived ambiguity with which the individual is confronted." Since individuals who are intolerant of ambiguity would theoretically have high attitudinal rigidity, those who are rigid should have low self-reliance. The highly anxious person would be expected to have high attitudinal rigidity. An unchanging self and environment might be seen by an anxious person as a relatively nonthreatening and nonanxious condition. The rigid individual would like to reduce the ambiguity and threat resulting from great differences between individuals. He therefore would reject others who are not like himself. Belief in and submission to authority might also serve to reduce ambiguity for the rigid person; thus appears the loading of authoritarianism on the rigidity factor [see Pyron (28)].

Simplicity-complexity of perceptual organization, tolerance-intolerance of ambiguity, and degree of attitudinal rigidity tend, it is suggested, to determine *S*'s capacity for experiencing stimuli beyond his familiar range. Simple or rigid *S*s, it is hypothesized, tend to reject stimuli beyond their familiar range while complex and flexible *S*s tolerate stimuli beyond their familiar range. In addition, *S*s who are perceptually complex and accept change may have experienced, in the past, a broader range of social, intellectual, or artistic stimuli. The concept of range of experience is similar to Kelman and Barclay's interpretation (12) of the F Scale as a measure of breadth of perspective.

In this study, measures of simplicity-complexity of perception, attitudinal rigidity, and the four variables associated with rigidity—self-reliance, anxiety, rejection of people and authoritarianism—were hypothesized to relate to acceptance of avant-garde art. It is suggested that much of avant-garde art tends to fall outside the familiar range of experience of most *S*s in a state college population. Those *S*s who are perceptually simple or rigid will tend to reject avant-garde art since they have, it is suggested, less capacity for accepting stimuli beyond their familiar range than *S*s who are perceptually complex or have low attitudinal rigidity. No measures of tolerance of ambiguity were given in this study. Intolerance of ambiguity is used only as a mediating concept to explore the relationships between simplicity-complexity or attitudinal rigidity and avant-garde art acceptance.

In much of traditional 20th century avant-garde art, it is suggested, there is relatively great ambiguity, complexity, novelty, and unexpected arrangement of artistic elements, while in popular and classical art there is a relatively great degree of order, familiarity, regularity, and predictability. In classical and popular representational painting, for example, familiar concrete human figures or objects are arranged in a predictable way on the canvas. Traditional painting devices such as realistic portrayal of figures and objects, shading, aerial and linear perspective, and familiar color are used to help insure that the perceiver's perceptual expectations will not be disconfirmed. Although *S*s may fail to perceive many of the subtle meanings in classical painting, they can interpret such painting in terms of the familiar objects represented. Avant-garde painting tends to disconfirm these conventional perceptual expectations. Unconventional and unexpected arrangement of forms make interpretation difficult in avant-garde painting. A comparison of classical paintings from

1500 to 1800 with 20th century Cubist, Surrealist, and Abstract Expressionist work will tend to confirm the assumption that there is greater ambiguity in avant-garde painting than in the earlier works.

The information contained in an avant-garde work may be interpreted by many perceivers as representing attitudes toward life or as ways of perceiving the world. The avant-garde movement, starting soon after 1900 in painting, literature, and music, was characterized by an attitude of freedom. The same attitude developed in drama, architecture, and sculpture [see Pyron (25, 26)]. In Cubist painting, Picasso and Braque violated many of the old rules of painting. The French poet Apollinaire practiced in his poetry the same freedom that his friend Picasso found in painting. Apollinaire freed his poetry from the rules of logical development, sentence structure, meter, and even of punctuation, which had made traditional poetry predictable and familiar. His lines are left open to the possibility of several combinations and interpretations. Much of later avant-garde poetry is based upon this freedom to rearrange the elements of the work. In the work of three other avant-garde poets used as examples in this study—that of Kenneth Patchen, Jean Arp, and Michael McClure—the same freedom is found.

The *S* high in need for simple order may reject avant-garde art because of its ambiguity. But he may also reject the attitude of freedom expressed in avant-garde work. New freedom gave the avant-garde artist opportunity to attack—sometimes in a subtle way and at other times in a blatant way—established attitudes and hypocrisies. Attacks in the areas of aggression and sexuality may be particularly threatening to the simple or rigid individual. The philosopher José Ortega y Gasset (21) has said that "The symbol of art is again seen in the magic flute of the Great God Pan which makes the young goats frisk at the edge of the grove" (21, p. 50). The simple and rigid person would rather keep the Pan within him deeply suppressed. Studies have strongly suggested that persons who overly incorporate the conventional norms and beliefs of middle-class culture and show concern with the expectations of others are more intolerant of ambiguity (3) and have greater attitudinal rigidity (27, 28) than persons who do not. *S*s who are perceptually simple or attitudinally rigid may reject avant-garde art because some of it threatens their conventional attitudes. The simple or rigid *S* would be likely to learn a generalized attitude of rejection of all avant-garde art. Almost all "modern" art might be seen by him as threatening or "immoral." An ambiguous painting, poem, or example of music might tend to elicit his generalized attitude of rejection.

Some of avant-garde art deals with irrational or unconscious processes. The principle of releasing the unconscious through improvisation was established by Surrealism in automatic writing and in dream-like pictures. Abstract Expressionist painting rather than Surrealist painting was used in this study. But the attitudes of Surrealist theory and practice are also found in Abstract Expressionist painting. The same attitudes are also found to some extent in the three avant-garde poets who came after Apollinaire—Arp, Patchen, and McClure. Jacob Korg (14) noted that ". . . Surrealist art is the result of this mingling of the commonplace and the fantastic, whose purpose is to override the usual limitations of perception . . ." (14, p. 461). Surrealism and some avant-garde poetry and painting which followed after its death in about 1940 employed the technique of remote association [see Mednick (19)] to evoke a spark of ambiguous but strongly felt meaning. The best of Surrealist art did, in the words of Shattuck (31) ". . . grasp the moment in its total significance, or more ambitiously, to manufacture a moment which surpasses our usual perception of time and space" (31, p. 266). Those who are rigid or prefer simple order, it is suggested, fear to "regress" to the preconscious or irrational thought processes and feelings. "Regression" would threaten to dissolve totally their established perceptual organization systems.

The American Abstract Expressionist painters realized more fully than the Surrealists the principle of automatic expression without censorship by rational rules. The Americans, however, were involved with a kind of motor manipulation of materials in a relatively spontaneous way. The European interest in specific associations joined together in a remote way was lost. The energized sloppy brush stroke became characteristic of the painting of Hofmann and De Kooning, whose work was used as examples in this study. An energized personal calligraphy characterizes the example of Pollock's work which was also used.

In the avant-garde music used in this study—that of Elgar Varése, Pierre Boulez, Karlheinz Stockhausen, and Lou Harrison—there is almost complete freedom from the rules of traditional or classical music. The music is either based to some extent upon an unfamiliar system, the 12-tone system, or is without recognizable rules of composition. Much of the music is not only atonal, but is without an easily perceptible melody. Boulez's "Musique Concréte," used in this study, was created by editing and rearranging taped sounds of real objects—e.g., the noises of trains—to produce a work of unconventional rhythmic structure. The electronic piece by Varése would be unfamiliar to *S*s who have not heard electronic music, not only because of its "incantory violence" [see Hodier (10)] but also because of its neglect of any conventions of harmony or melody. Harrison's piece has a melody line running through it, but depends upon the arrangement of percussion rhythms obtained by unusual instruments. These four works were selected as representative of avant-garde music, since they are clearly different from classical music. Other examples of 20th century music might not seem so different from classical music by musically untrained *S*s. Mikol (20) found no difference between dogmatic and nondogmatic *S*s in their acceptance of the music of Bartok, but did find differences between the two groups for Schoenberg's music, which is based upon a radically different system of composition.

Simple or rigid *S*s will, it was hypothesized, reject these musical examples because the music is unfamiliar. They will prefer popular or classical music. For the majority of *S*s, the music of Beethoven would be more acceptable than avant-garde music. Beethoven's music belongs to a class which is familiar to them. They are able to organize the elements of the music as it unfolds in time into a unitary pattern, since it is based upon a familiar system of composition. An *S* familiar with "rock and roll" music could predict even more effectively the unfolding of the musical elements in a piece such as Elvis Presley's "Beach Boy Blues." Its rhythm and words repeat with a highly certain regularity.

B. ATTITUDINAL RIGIDITY AND SIMPLICITY-COMPLEXITY OF SOCIAL ORDERING

The variables which loaded above .30 [see Pyron (28)] on the following two factors were hypothesized to relate to acceptance of avant-garde art.

1. *Attitudinal Rigidity: Rejection-Acceptance of Attitude Positions and Social Stimuli Tending to Threaten the Perceptual Ordering and Belief Systems*

a. Acceptance-rejection of change as an overriding attitude. The central hypothesis is that the *S* who accepts change is open to new experiences that might lead to a change in his perceptual organization and belief systems. The *S* who rejects change is closed to new experiences that might threaten his rigidly held attitudes and beliefs. The *S* who rejects change is hypothesized to reject novel and new combinations of artistic elements. It is therefore hypothesized that *S*s high in acceptance of change will accept avant-garde art significantly more than those low in acceptance of change.

b. Authoritarianism. Since the F Scale [see Adorno, *et al.* (1)] attempts to measure conventionalism and rigidity as well as authoritarianism, it is hypothesized that those high on the F Scale will reject novel and new combinations of artistic elements. In addition, the content of avant-garde art might threaten the authoritarian's highly conventionalized perceptual organization and belief systems. *S*s who score low on the F Scale (nonauthoritarian) will accept avant-garde art significantly more than those who score high (authoritarian) on the F Scale.

c. Rejection of people who are dissimilar to one's self. The rigid *S*, it is suggested, rejects others who are not like himself, since he does not tolerate ambiguity resulting from great differences between people. He maintains perceptual and emotional distance from these others through rejection and perceptual simplification. It is hypothesized that the attitude of rejection of stimulus objects that are beyond the *S*'s accepted range of stimuli is generalized; rejection of stimuli beyond the familiar range applies to both people and art. *S*s who reject others will also reject new and novel combinations of artistic stimuli and the content of avant-garde art. It is hypothesized that *S*s low in rejection of people will accept avant-garde art significantly more than those high in rejection of people.

d. Reliance upon self vs. reliance upon others. Those who are extremely concerned with other people's evaluations of themselves are, it is suggested, also extreme in their incorporation of the conventional organization scales and attitude positions of the culture. Conformity to the values of others tends to reduce the threat of ambiguity. It is hypothesized that the *S* who is not self-reliant cannot creatively participate in novel situations or accept novel stimuli because of his intolerance of ambiguity. The *S* who is extremely concerned with the evaluations of others is often afraid to experience works of art involving new combinations of artistic elements, since he fears others might think him different or unconventional. It is hypothesized that *S*s high in self-reliance will accept avant-garde art significantly more than those low in self-reliance.

e. Level of manifest anxiety. Those who are highly anxious often desire to maintain their attitudes, self-concepts, and perceptual organization systems in an unchanging condition. Stimuli that might threaten to change their mechanisms for relating to themselves and the world are avoided by the highly anxious person. In addition, the highly anxious person may expend so much energy in being anxious that he is unable to respond emotionally or perceptually to novel or ambiguous stimuli. It is hypothesized that those low in manifest anxiety will accept avant-garde art significantly more than those high in anxiety.

2. *Simplicity-Complexity of Social Ordering*

Two constructs dealing with the degree of ordinal perceptual simplification have been discussed: (*a*) number of discriminable categories along scales, and (*b*) number of different scales available to order stimuli.

a. Sexual restraint vs. affirmation of sexuality. A measure of sex attitude loaded highest on the simplicity-complexity of social ordering factor [see Pyron (28)]. This was the only attitudinal variable which was part of this factor. The hypothesis is that individuals who make a strong separation between the sexual and the nonsexual in heterosexual relations, whether they exhibit a negation or an affirmation of the sexual, are more extreme in imposing simple, abstract, and dichotomous order upon other aspects of social behavior than individuals who do not make such a separation. Sexual liberty has often been associated in popular thinking with 19th and 20th century avant-garde art movements. If sexual freedom is part of a general attitude of freedom, toleration of ambiguity, and ability to experience emotion, it would follow that individuals who endorse sexual freedom would also endorse avant-garde art significantly more than those who prefer sexual restraint.

b. Number of scales of discrimination. To achieve simplicity of ordinal perceptual organization, the *S* uses a *very small* number of scales to order the behavior of others. An *S* who uses a small number of scales (simple organizer) is theoretically less able to discriminate among social stimuli than an *S* that uses a large number of scales, while the *S* who uses a large number of scales (complex organizer) makes greater differentiation between social stimuli than the simple organizer *S*. The complex organizer not only tolerates greater complexity in his environment than does the simple organizer, but he perceives his environment as being more complex than does the simple organizer. Artistic communication at its best, it is suggested, transcends the conventional simplified mechanisms of perception. Much of avant-garde art cannot be easily perceived and interpreted by use of simple conventional ordering scales. The complex organizer is likely to prefer artistic stimuli that are not easily ordered on conventional scales. He is also likely to have the ability to achieve a more complex ordering of ambiguous stimuli than the simple organizer. It is hypothesized that *S*s who utilize a large number of scales of discrimination will accept avant-garde art significantly more than those who utilize a small number of scales.

c. Number of discriminable categories along scales. The research of Volkmann (34) and others suggests that the number of discriminable categories along scales in ordinal perceptual organization is determined by the range of stimuli the organism has experienced in the past and is experiencing at the time of judgment. An overriding attitude position affirming a dichotomous type of organization would tend to reduce the number of discriminable categories *S* uses. This attitude position is learned from the culture, but to different degrees by different individuals. The traditional Western way of abstracting and experiencing is dichotomous—i.e., an object is either A or Not-A; a person is either good or bad, saved or not saved; and there are no discriminable points in between. The *S* who experiences social or art stimuli on the basis of a very small number of categories ignores a great deal of sensory information. He tends to lose perceptual and emotional contact with stimuli. Dichotomizing is a way of maintaining perceptual and emotional distance from threatening objects—people or art. Ambiguous art would thus tend to threaten *S*s who tend to dichotomize, while it would tend to be preferred by those who make use of a large number of discriminable categories. It is hypothesized that those *S*s who are high in number of categories will prefer avant-garde art significantly more than those low in number of categories.

C. METHOD

1. *Design*

All *S*s were given examples of three art esthetics: (a) popular, (b) classical, and (c) avant-garde. There were three art mediums within each esthetic: (*a*) painting, (*b*) literature, and (*c*) music. The term "esthetic" is used here only for convenience in classification and it is not implied, for example, that the esthetic of avant-garde music is identical to the esthetic of avant-garde poetry, but that avant-garde music and poetry both involve a freedom from traditional rules of composition.

2. *Procedure*

*S*s in groups of three to five were asked to evaluate each of four examples of art for all three esthetics and all three mediums. There were thus 36 selections of art. Within each medium the 12 examples were presented to the *S*s in four randomly determined orders. The *S*s read literature examples first, then they were shown examples of painting, and finally *S*s listened to tape recordings of music.

*S*s were required to evaluate each of the 36 works on a Semantic Differential [see Osgood, Tannenbaum, and Suci (22)] consisting of 12 bipolar adjectives. Eight of these adjectives were reported by Osgood, *et al.* to have loadings of .7 or higher on the evaluative factor. These eight bipolar adjectives were 1) good-bad, 2) honest-dishonest, 3) nice-awful, 4) sacred-profane, 5) kind-cruel, 6) valuable-worthless, 7) clean-dirty, and 8) happy-unhappy. The four additional bipolar scales were, 9) emotional-unemotional, 10) creative-uncreative, 11) interesting-uninteresting, and 12) inspiring-uninspiring. An evaluative score for each of the 36 works for each *S* was obtained by summing the 12 positions checked along the nine-point scales and dividing by 12. Then, each *S*'s evaluative scores were averaged across the four selections within each combination of esthetic type and medium. For example, evaluative scores for avant-garde painting could range from one to nine. The higher the evaluative score, the greater the acceptance of the esthetic-medium combination.

3. *Art Selections*

a. Popular art. Painting: color reproductions were shown of (*a*) George Hinke's *Landscape,* (*b*) Cid Kreisbach's *Winter Scene,* (*c*) Francis Hook's *Footprints in the Snow,* and (*d*) J. Greenleaf's *Landscape.* Literature: selections of approximately 100 words each were taken from (*a*) Ayn Rand's *Fountainhead,* (*b*) Rachel Carson's *Silent Spring,* (*c*) Jim Bishop's *The Day Lincoln Was Shot,* and (*d*) Harper Lee's *To Kill a Mockingbird.* Music: one minute excerpts were taken from (*a*) the Beatles' "All My Loving," (*b*) Dion and the Belmont's "Lovers Who Wander," (*c*) Elvis Presley's "Beach Boy Blues," and (*d*) Buddy Holly's "Bo Diddly."

b. Classical art. Painting: color reproductions were shown of (*a*) da Vinci's *Mona Lisa,* (*b*) El Greco's *Toledo in a Storm,* (*c*) Rembrandt's *The Syndicies of the Drapers Guild,* and (*d*) Dürer's *Portrait of Hieronymus Holzschuher.* Literature: selections of 100 words each were taken from (*a*) Milton's *Areopagitica,* (*b*) Cervantes' *Don Quixote,* (*c*) Shakespeare's *Hamlet,* and (*d*) Edmund Spenser's *The Faerie Queen.* Music: one minute excerpts were taken from (*a*) Chopin's "Etude, Op. 25 in A Flat," (*b*) Beethoven's *Ninth Symphony,* (*c*) Liszt's "La Campanella," and (*d*) Beethoven's "Cello Sonata, Op. 5, No. 1."

c. Avant-garde art. Painting: color reproductions were shown of (*a*) Jackson Pollock's *Number 12,* (*b*) de Kooning's *Composition,* (*c*) Hofmann's *Exuberance,* and (*d*) Tomlin's *In Praise of Gertrude Stein.* Literature: selections of 100 words each were taken from (*a*) Jean Arp's seminonsense automatic writing from his *On My Way,* (*b*) Apollinaire's "Aim" from his *Calligrammes,* (*c*) Kenneth Patchen's *Sleepers Awake,* and (*d*) Michael McClure's *The New Book: A Book of Torture.* Music: one minute excerpts were taken from (*a*) Varése' "Poem Electronique," (*b*) Stockhausen's "No. 5 Zeitmasse," (*c*) Lou Harrison's "Canticle No. 3," and (*d*) Boulez's "Musique Concréte."

4. *Subjects*

Forty-eight *S*s, 14 females and 34 males, from an introductory-psychology class at Wisconsin State University, Whitewater, were first given the nine personality measures and then the art evaluation task.

5. *The Nine Personality Measures*

These measures were identical to the measures given in the previous factor-analytic study [see Pyron (28)].

a. Attitudinal rigidity factor.

(1). *Openness vs. closedness of beliefs—The Dogmatism Scale.* Developed by Rokeach (29), this scale is construed as a measure of openness *vs.* closedness of belief systems. It is assumed to be independent of the content or direction of belief, and measures openness to new concepts that might change the belief system *vs.* closedness to concepts and situations that might threaten one's established belief system. The 40-item Form E of the Dogmatism Scale was used, and *S*s were provided with a six-point scale with six verbal anchors for each statement. Agreement with any statement was scored in the dogmatic direction. The Dogmatism Scale loaded .56 on the attitudinal rigidity factor.

(2). *Acceptance of change—the Change Inventory.* The items of the Change Inventory [see Pyron (27, 28)] fall into the following groups: (*a*) a desire for order, certainty, and a knowledge of what one will be doing in the future, (*b*) the value of flexibility, (*c*) acceptance of change in human society, (*d*) sensitivity to change in the natural world, (*e*) a desire to live for the present rather than the future, and (*f*) affirmation of new experiences which change one's outlook. Ten of the 18 items of the Change Inventory were taken from the Belief Q-sort [see Pyron (24)], and new items were created on the basis of test-consistency item analyses of the Belief Q-sort. To control for possible yea-saying and nay-saying response sets, *S*s were required to agree with exactly half of the statements and disagree with the other half. Half of the items were worded in a positive direction and half in a negative direction. Agreement with a positive statement was scored and disagreement with a negative statement was scored. Each statement was rated along an 18-point scale with six verbal anchors: the higher the score the greater the acceptance of change. The Change Inventory loaded —.53 on the attitudinal rigidity factor.

(3). *Authoritarianism—the F Scale.* The 28-item version of the original F Scale [see Adorno, *et al.* (1)] with six scale positions and six verbal anchors for each statement was used: the higher the score the greater the authoritarianism. The F Scale loaded .50 on the attitudinal rigidity factor.

(4). *The rejection of people test.* Eight statements dealing with prejudice toward Jews, Negroes, and other minority groups were taken from the

Ethnocentricism Scale of Adorno, *et al.* Seven new statements were written [see Pyron (28)] that expressed rejection of "Beats," the Student Socialist Party, "arty" students, students who have wild parties, and men with long hair or beards. *S*s responded to each item on a six-point scale: the higher the score the greater the rejection of people. The rejection of people test loaded .45 on the attitudinal rigidity factor.

(5). *The self-reliance test.* The concept of self-reliance concerns the extent to which an individual relies upon his own judgments, his own experiences, and his own scale positions instead of depending upon those of others. Six of the 14 items of this scale were taken from the Belief Q-sort (24), and the new items were originated on the basis of test-consistency item analyses of the Belief Q-sort. The items fall into the following three groups: (*a*) the desire to conform to the expectations of others regarding behavior, (*b*) the desire to lose oneself in group activity *vs.* the desire to be alone, and (*c*) the desire to be an individual, to develop one's values and attitudes on the basis of one's own experiences rather than on the basis of those of others. Half of the items were worded in a positive direction (self-reliant) and half in a negative direction (reliance upon others). Agreement with a positive statement was scored and disagreement with a negative statement was scored. Each statement was rated along an 18-point scale with six verbal anchors: the higher the score the greater the self-reliance. The self-reliance test loaded —.37 on the attitudinal rigidity factor.

(6). *Taylor Manifest Anxiety Scale.* The 50-item form of the Taylor Scale (32) was used and each item was rated on a six-point scale with six verbal anchors: the higher the score the greater the manifest anxiety. The Taylor Scale loaded .36 on the attitudinal rigidity factor.

b. Simplicity-complexity of social ordering factor

(1). *Sex attitude—the Abstract Sex Inventory.* The general concept concerns the acceptance of sexuality and the desire to experience sex fully *vs.* the desire to restrain one's sexual impulses. The four items of this scale are stated abstractly: for example, "It is against my own personal beliefs to have sexual intercourse before marriage." Each item was rated on an 18-point scale with six verbal anchors: the higher the score the more affirmative the attitude toward sexuality. In a previous pilot study, the test-retest reliability for this test was .62 for females and .79 for males (University of Wisconsin at Madison sample). The Abstract Sex Inventory loaded —.45 on the simplicity-complexity of social ordering factor.

(2). *Number of categories along scales—the Adjective Scale.* *S*s were asked to fit real people known to them to eight roles, such as "The most outstanding teacher you have known." Each of the eight role persons was rated on a 30-item adjective list taken largely from the Semantic Differential [see Osgood, *et al.* (22)]. Nine scale positions and four verbal anchors were provided for each of the 30 adjective scales. To the extent that *S* gave different scale positions to different role persons, it was assumed that he was able to discriminate among them along the scales. The Adjective Scale was scored by counting the number of different scale positions used for each adjective scale, summed across all 30 scales: the higher the score the greater the number of categories along the scales. The Adjective Scale loaded —.39 on the simplicity-complexity of social ordering factor.

(3). *Number of scales used to discriminate between persons—The Revised Role Repertory Test.* This test was developed from Kelly's original Role Construct Repertory Test (11). *S*s were again asked to fit real people to seven roles, six being the same as in the Adjective Scale. Then they considered these role persons in pairs. Twenty-one pairs were used in this study. *S*s were told they must give five characteristics for each pair of individuals, ways in which the two were alike or different. The measure was the number of different bipolar scales used by *S* to describe the ways the members of a pair differed (scales used in describing how members were alike were not scored): the higher the score the greater the number of scales of discrimination used to differentiate the persons. Bieri (2) reports a reliability of .80 (Kuder-Richardson 21) for the original Role Construct Repertory Test. The Revised Role Repertory Test loaded —.35 on the simplicity-complexity of social ordering factor.

D. RESULTS AND DISCUSSION

1. *Overall Effects of Art Medium and Art Esthetic*

A three by three repeated measures analysis of variance indicated that both art medium and art esthetic were significant beyond the .01 level. Esthetic was a more powerful variable ($F = 102.8$) than medium ($F = 9.3$).

Painting was preferred (mean evaluation = 5.8) to literature (mean = 5.5) and music (mean = 5.4). Duncan Range Tests (6) indicated that painting was liked significantly more than both literature and music, at the .01 level, but literature and music did not differ significantly.

Popular art was given first preference (mean = 6.1, on the nine-point scale), classical second (mean = 5.8), and avant-garde was liked least of all (mean = 4.7). All three esthetics differed significantly at the .05 level and beyond when the means were tested by the Duncan Range Test.

The significant interaction between art medium and art esthetic (at the .01 level) indicates that although painting was given first preference within the popular and classical esthetics, this was not the case within the avant-garde esthetic. Within the avant-garde esthetic, literature was given first preference. This finding might tentatively suggest that intolerance of ambiguity in art-meaning is a more important factor than rejection of the attitude content of the art. Abstract Expressionist painting has little explicit attitude content, while the avant-garde literature used in this study has more attitude content. But there are other possible interpretations of this interaction. Since *S*s have seen more examples of avant-garde painting in popular media than examples of avant-garde poetry, an attitude of rejection of "modern" painting on the basis of both ambiguity and "immoral" attitude content might generalize to the four examples of Abstract Expressionist painting. Since *S*s are not familiar with the type of avant-garde poetry used in this study, they have not developed a generalized attitude of rejection toward this art.

2. *Effects of the Nine Personality Measures*

The hypothesis was made that *S*s who score high on attitudinal rigidity and on its associated variables and high on simplicity of perceptual ordering would reject avant-garde art significantly more than those who score low. It is also of interest to determine the relationship between the nine personality measures and acceptance-rejection of popular and classical art.

To test the hypotheses, the entire group of 48 *S*s was divided into four groups on each personality variable, with 12 *S*s in each of the four groups. Nine separate four by three by three repeated measures analyses of variance were done. A usual within-sums-of-squares error term (error *a*) was used to test the overall effects of each personality variable, and a pooled-interactions term (error *b*) was used to test the interaction between each personality variable—taken one at a time—and art esthetic and art medium. Duncan Range Tests were used to test the significance of the differences between means. When the means being compared were based upon the scores of the same *S*s, error *b* was used, and when the means were based on different *S*s, error *a* was used in computing the significant studentized ranges. When a personality variable failed to reach overall significance as a four category variable, it was tested by dichotomizing *S*s into highs and lows.

For five of nine personality measures hypothesized to relate to the acceptance of avant-garde art, the initial hypotheses were confirmed. For these five measures, significant overall effects were significant. Nine of the interactions with art esthetic or art medium was significant. But only two of these measures—acceptance of change and number of scales of discrimination—were significant as four category variables by analyses of variance. The other three measures—rejection of people, self-reliance, and dogmatism—were significant only as dichotomous variables. For all five measures which were significant by analyses of variance, the differences between the two extremes, highs and lows, were significant by Duncan Range Tests. The significant differences reported in Tables 1 and 2 are based on Duncan Range Tests. In all cases only the highs and lows were significantly different. None of the personality measures was significantly related to the acceptance of popular and classical art. Those most extreme in dogmatism, however, liked popular art significantly more than classical art.

TABLE 1
MEAN EVALUATION SCORES OF FOUR VARIABLES OF FACTOR A FOR POPULAR (P), CLASSICAL (C), AND AVANT-GARDE (A-G) ART

Individual difference variable	P art	C art	A-G art	P *vs.* C	P *vs.* A-G	C *vs.* A-G
Dogmatism (Overall A-G $F = 5.44^*$—as a dichotomous variable)						
Lows (least dogmatic)	5.7	5.8	4.9	.1	.8**	.9***
Low mediums	6.4	6.1	5.1	.3	1.3†	1.0***
High mediums	6.1	6.0	4.7	.1	1.4†	1.3†
Highs (most dogmatic)	6.2	5.6	4.2	.6*	2.0†	1.4†
Difference between highs and lows	.5	.2	.7*			
Acceptance of change (Overall A-G $F = 3.70^*$)						
Lows (most rejecting)	6.3	5.9	4.4	.4	1.9†	1.5†
Low mediums	5.9	5.8	4.6	.1	1.3†	1.2†
High mediums	6.1	5.9	4.6	.2	1.5†	1.3†
Highs (most accepting)	6.0	6.0	5.4	.0	.6*	.6*
Difference between highs and lows	.3	.1	1.0**			
Rejection of people (Overall A-G $F = 5.53^*$—dichotomous)						
Lows (least rejecting)	5.9	6.0	5.1	.1	.8**	.9***
Low mediums	6.0	5.8	5.0	.2	1.0†	.8**
High mediums	6.1	5.5	4.6	.6	1.5†	.9***
Highs (most rejecting)	6.3	6.3	4.3	.0	2.0†	2.0†
Difference between highs and lows	.4	.3	.8*			
Reliance upon self (Overall A-G $F = 8.38^{**}$)						
Lows (least reliant)	6.2	5.9	4.4	.3	1.8†	1.5†
Low mediums	6.1	5.7	4.5	.4	1.6†	1.2†
High mediums	6.0	5.9	5.0	.1	1.0***	.9***
Highs (most reliant)	6.1	6.1	5.1	.0	1.0***	1.0†
Difference between highs and lows	.1	.2	.7*			

* $p = .05$.
** $p = .01$.
*** $p = .005$.
† $p = .001$.

TABLE 2
MEAN EVALUATION SCORES OF ONE VARIABLE OF FACTOR B, NUMBER OF D SCALES, FOR POPULAR (P), CLASSICAL (C), AND AVANT-GARDE (A-G) ART

Individual difference variable	P art	C art	A-G art	P *vs.* C	P *vs.* A-G	C *vs.* A-G
Number of D Scales (Overall A-G $F = 2.89$*)						
Lows (most simple)	6.2	6.1	4.1	.1	2.1†	2.0†
Low mediums	5.9	5.9	5.1	.0	.8**	.8**
High mediums	6.2	5.9	5.0	.3	1.2†	.9***
Highs (most complex)	6.0	5.8	5.0	.2	1.0***	.8**
Difference between highs and lows	.2	.3	.9*			

* $p = .05$.
** $p = .05$.
*** $p = .005$.
† $p = .001$.

a. Attitudinal rigidity factor. Four of the six variables of this factor reached overall significance.

(1). *Dogmatism.* *S*s who scored in the lowest 50 per cent on dogmatism (least dogmatic) accepted avant-garde art significantly ($F = 5.44$, $p < .05$) more than those who scored in the highest 50 per cent on dogmatism (most dogmatic). Table 1 also indicates that for the most dogmatic *S*s the differences between the avant-garde and both popular and classical means were 2.0 and 1.4 respectively, while for the *S*s low in dogmatism these differences were only .8 and .9. *S*s who tend toward openness in their belief systems like avant-garde art almost as well as popular and classical art. Those who are most closed in their belief systems like popular art significantly more than classical art (at the .05 level), and like popular art extremely more than avant-garde art (significant at the .001 level).

(2). *Acceptance of change.* Acceptance of change was significantly related to the acceptance of avant-garde art as a four category variable ($F = 3.70$, $p < .05$). Table 1 shows that those high in acceptance of change liked avant-garde art significantly more than those low in acceptance of change. However, the difference between acceptance of popular and classical art for the group most extreme in rejection of change was not significant. Those who were highest in acceptance of change liked avant-garde art almost as well as popular and classical art (the mean differences were only .6 and .6 respectively), while those most extreme in rejecting change liked popular art considerably more than avant-garde (the mean difference in liking was 1.9, significant at the .001 level).

(3). *Rejection of people.* *S*s who scored in the lowest 50 per cent in rejection of people accepted avant-garde art significantly ($F = 5.25$, $p < .05$)

more than those who scored in the highest 50 per cent in rejection of people. Table 1 shows that those who were least rejecting of people liked avant-garde art almost as well as popular and classical, but those most rejecting of people liked popular and classical art considerably more than avant-garde art (the mean differences were 2.0 in both cases).

(4). *Reliance upon self.* *S*s who scored in the highest 50 per cent on self-reliance liked avant-garde art significantly ($F = 8.38$, $p < .01$) more than *S*s who scored in the lowest 50 per cent on self-reliance. In Table 1, the differences in acceptance of avant-garde art between the extreme highs and lows were the only important differences for this measure.

b. Simplicity-complexity of social ordering factor. Number of scales of discrimination: within the simplicity-complexity factor, only number of scales of discrimination (number of D Scales), measured by the Revised Role Repertory Test, was significantly related to avant-garde art acceptance. In addition, none of the variables of this factor was significantly related to acceptance of popular or classical art. An analysis of variance showed number of D Scales to be significantly related to avant-garde art acceptance as a four category variable ($F = 2.89$, $p < .05$). Table 2 also shows that the group of *S*s who used the largest number of different scales to discriminate between role persons (complex organizers) accepted avant-garde art significantly more than the group who used the smallest number of scales (simple organizers). No other differences in Table 2 were of importance.

3. *Significance of the Results*

The data do tend to support the initial hypothesis that those who are highest in attitudinal rigidity and highest in simplicity of perceptual organization would reject avant-garde art more than those who accept change and tolerate a wider range of social stimuli and are complex in perceptual organization.

The possible significance of this study lies in its extension of concepts of simplicity-complexity of perception, tolerance-intolerance of ambiguity, and attitudinal rigidity to art acceptance. Those who tend to be simple in perceptual organization, intolerant of ambiguity, and who desire to maintain themselves and their world in an unchanging condition, it has been suggested, tend to reject stimuli that are unfamiliar, difficult to interpret, or threatening to their attitudes. An art object, a poem, or a musical composition can be construed as being a configuration of stimuli. Rigid and perceptually simple *S*s would also be expected to reject unfamiliar or ambiguous concepts, machines, or even unfamiliar plants.

One theory of perceptual organization [see Kelly (11) and Pyron (27, 28)] assumes that scales for organizing sensory input of all kinds are hierarchically organized, with scales for registering sensory information at the bottom and the most global, integrating scales like good-bad at the top. These master integrating scales at the top are commonly called belief or attitude scales, with indifference points (9) determined by the range of concrete stimuli the *S* has experienced and by other overriding scales. In the present study not only the degree of complexity of the perceptual system itself, but also the positions on several abstract master scales investigated were related to the acceptance of avant-garde art. The kinds of stimuli the *S* enjoys and accepts are determined by his position on these master scales.

E. SUMMARY

Forty-eight college *S*s indicated the degree to which they accepted three art esthetics, (*a*) popular, (*b*) classical, and (*c*) avant-garde; and three art mediums, (*a*) painting, (*b*) literature, and (*c*) music. There were thus nine esthetic-medium combinations, and for each combination four examples were given, making a total of 36 art selections. All *S*s evaluated all 36 selections. It was suggested that in avant-garde art there is greater ambiguity, novelty, and unexpected arrangement of elements than in popular and classical art, while in the latter there is a greater degree of order, regularity, and predictability than in the former. Art evaluation was studied as a function of two personality factors, (a) attitudinal rigidity and (b) simplicity-complexity of social ordering. Individual differences in preference for avant-garde art were significantly related to the degree of complexity of the perceptual ordering system and to positions on four of six variables of the rigidity factor. These four variables were, (*a*) dogmatism, (*b*) acceptance of change, (*c*) rejection of people, and (*d*) reliance upon self.

REFERENCES

1. ADORNO, T., FRENKEL-BRUNSWIK, E., LEVINSON, D., & SANFORD, R. N. The Authoritarian Personality. New York: Harper, 1950.
2. BIERI, J. Cognitive complexity and predictive behavior. *J. Abn. & Soc. Psychol.*, 1955, **51**, 263-268.
3. BRUNER, J. S., GOODNOW, J. J., & AUSTIN, G. A. A Study of Thinking. New York: Wiley, 1956.
4. BUDNER, S. Intolerance of ambiguity as a personality variable. *J. Personal.*, 1962, **30**, 29-50.
5. DE SOTO, C. B. Two-category judgments of sequences of stimuli of two values. *J. Exper. Psychol.*, 1958, **55**, 34-38.
6. DUNCAN, D. B. Multiple range and multiple F tests. *Biometrics*, 1955, **11**, 1-42.

7. FRENKEL-BRUNSWIK, E. Intolerance of ambiguity as an emotional and perceptual personality variable. *J. Personal.*, 1949, **18**, 108-143.

8. HARVEY, O. J., HUNT, D. E., & SCHRODER, H. M. Conceptual Systems and Personality Organization. New York: Wiley, 1961.

9. HELSON, H. Adaptation-Level Theory. New York: Harper & Row, 1964.

10. HODIER, A. Since Debussy: A View of Contemporary Music. New York: Grove Press, 1961.

11. KELLY, G. A. The Psychology of Personal Constructs (Vol. I). New York: Norton, 1955.

12. KELMAN, H. C., & BARCLAY, J. The F Scale as a measure of breadth of perspective. *J. Abn. & Soc. Psychol.*, 1963, **67**, 608-615.

13. KENDIG, I. Studies in perseveration: II. Determining factors in the development of compulsive activity. *J. of Psychol.*, 1937, **3**, 231.

14. KORG, J. Modern art techniques in "The Waste Land." *J. Aesthet. & Art Criticism*, June, 1960, XVIII (4), 456-463.

15. KOUNIN, J. Experimental studies of rigidity. *Charac. & Personal.*, 1941, **9**, 251.

16. LEVENTHAL, M. Cognitive processes and interpersonal prediction. *J. Abn. & Soc. Psychol.*, 1957, **55**, 176-180.

17. LUNDY, R. M., & BERKOWITZ, L. Cognitive complexity and assimilative projection in attitude change. *J. Abn. & Soc. Psychol.*, 1957, **55**, 34-37.

18. MAYO, C., & CROCKETT, W. H. Cognitive complexity and primacy-recency effects in impression formation. *J. Abn. & Soc. Psychol.*, 1964, **68**, 335-338.

19. MEDNICK, S. A. The associative basis of the creative process. *Psychol. Rev.*, 1962, **69**, 220-232.

20. MIKOL, B. The enjoyment of new musical systems. In M. Rokeach, *The Open and Closed Mind*. New York: Basic Books, 1960. Chap. 15.

21. ORTEGA Y GASSET, J. The Dehumanization of Art. Princeton, N.J.: Princeton Univ. Press, 1948.

22. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. The Measurement of Meaning. Urbana: Univ. Illinois Press, 1957.

23. PALLY, S. Cognitive rigidity as a function of threat. *J. Personal.*, 1955, **23**, 346-355.

24. PYRON, B. An attempt to test the theory of psychological development. *Psychol. Rep.*, 1959, **5**, 685-698.

25. ———. Wright's diamond module houses: His development of non-rectilinear interior space. *Art Journal*, Winter 1961-1962, **21**, 92-96.

26. ———. The Tao and Dada of recent American ceramic art. *Artforum*, March 1964, **11**, 41-44.

27. ———. Intercorrelations of simplicity-complexity of social ordering, accuracy of perception, acceptance of change, self-reliance and sex attitude. *Percept. & Motor Skills*, 1965, **20**, 1293-1304.

28. ———. A factor-analytic study of simplicity-complexity of social ordering. *Percept. & Motor Skills*, 1966, **22**, 259-272.

29. ROKEACH, M. The Open and Closed Mind. New York: Basic Books, 1960.

30. SCOTT, W. A. Cognitive complexity and cognitive flexibility. *Sociometry*, 1962, **25**, 405-414.

31. SHATTUCK, R. The Banquet Years: The Arts in France, 1885-1918. New York: Harcourt Brace, 1958.

32. TAYLOR, J. A. A personality scale for manifest anxiety. *J. Abn. & Soc. Psychol.*, 1953, **48**, 285-290.

33. VANNOY, J. S. Generality of cognitive complexity-simplicity as a personality construct. *J. Personal. & Soc. Psychol.*, 1965, **2**, 385-396.

34. Volkmann, J. The anchoring of absolute scales. *Psychol. Bull.*, 1936, **33**, 742-743.

35. Zajonc, R. B. The process of cognitive tuning in communication. *J. Abn. & Soc. Psychol.*, 1960, **61**, 159-167.

Instructional Research Laboratory
The University of Wisconsin
202 State Street
Madison, Wisconsin 53706

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 179-195.

RUNWAY PERFORMANCE AS A FUNCTION OF HUNGER LEVEL MEASURED AS FOOD ACCEPTANCE*[1]

Tufts University

WARREN H. TEICHNER

A. INTRODUCTION

This experiment was part of a larger investigation concerned with rodent repellents. One part of that investigation involved mixing chemicals known to be repellent into the animal's daily food ration and evaluating repellency in terms of food consumption. A long-term feeding study demonstrated that, after an initial rejection, rats on a constant food deprivation schedule adjust their food intake as a function of the concentration of the chemical [see Teichner, Rowntree, and Wagner (6)]. This result is of interest as a motivational phenomenon because it offers a novel operational definition of level of hunger in terms of amount of food acceptance.

Conventional procedures for defining hunger drive are based on the relationship between hours of food deprivation and one or another of food consumption, body weight, or, more recently, readiness to eat [see Bolles (1, 2)]. The thinking underlying the use of these definitions is that if a reliable relationship can be established, hours of deprivation can be used as an index of hunger level in studies of learned performance. While such thinking is reasonable in principle, few, if any, reliable relationships have actually been established. An exception may be the dependence of body weight on hours of deprivation, at least for times longer than 12 hours (2). Nevertheless, neither body weight nor its correlate, readiness to eat, has been shown to be related systematically to a measure of behavioral interest. In fact, as Bolles notes, Davenport and Goulet (3) have demonstrated a gradually decreasing performance for animals maintained on a constant body weight.

One likely difficulty with deprivation time definitions is that the duration

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 11, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright, 1966, by The Journal Press.

[1] The experimental work and data analysis were done at the University of Massachusetts under Contract DA19-129-AMC-236 (N) with the U. S. Army Natick Laboratories. This paper was prepared at Tufts University under Contract DA-120-AMC 0691 (N) with the U. S. Army Natick Laboratories. Thomas Fonda and Terri Soward assisted in collecting and handling the data. Richard Christ provided programming and computational assistance. The U. S. Fish and Wildlife Service, Denver, Colorado, furnished the chemical repellent.

of time without food is confounded with events which are a function of that time. For example, activity and water consumption vary diurnally and with the length of the deprivation cycle. In turn, these events influence body weight, food consumption, etc. As an index of hunger level, time without food does not describe positions along a dimension. Instead, it codes the differences between different deprivation times due to uncontrolled activities that intervene between removal of the food and the time of food presentation.

A related methodological difficulty concerns the logic of operational definition. Experiments that attempt to relate learned response strength to hunger are necessarily performed before the animals are fed and, therefore, the experiment itself is among the events intervening between the initiation and the end of deprivation. As such, it is part of the definition of hunger rather than independent of it as is desired. The objection is trivial unless the learned performance actually affects the hunger level—for example, by affecting food consumption. The data to be reported have a bearing on this question.

The ideal definition of hunger level is one based on demonstrable effects occurring just *prior* to the measure of performance. The hunger measure may then be treated as an effect that is dependent upon a manipulable condition. It should also be an effect which is unambiguously related to that condition. These requirements appear to be met by using amount of food accepted (consumed) as a function of nontoxic repellent concentration. This kind of definition allows the experimenter to demonstrate the actual level of hunger before every later measure of response strength. Very simply, for a constant deprivation time, different levels of hunger are defined by the degree of subject-controlled satiation. Satiation or food acceptance is inversely related to concentration of repellent: the more food accepted, the lower the hunger drive. This approach not only permits a prior definition of hunger, but allows for the relation of individual response strengths to prior individual hunger levels.

It is to be expected that, within limits for a constant repellent concentration, as deprivation time increases the amount of repellent food eaten will also increase. This presents no logical problem to an acceptance concept of hunger. It means only that time of deprivation, and the confounding factors which occur within it, are factors on which hunger level depends rather than factors with which it is defined. The dependence is a matter for experimental analysis; the definition of hunger is independent of this analysis and can be applied without waiting for it.

It has already been shown that food acceptance varies with concentration of repellent (6). It still remains to be shown that learned performance varies with level of hunger defined this way. This was the purpose of the present experiment. Study was made of the interrelations among runway performance measures, food acceptance prior to runway exposure, food consumption in the goal box following running, and body weight. In addition, the amount of physical work required to reach the goal box was varied as a means of varying the sensitivity of the performance measurement to the effect of hunger level.

B. METHODS

1. *Apparatus*

Fifteen individual rat-feeding chambers, 8 3/8 × 4 1/2 × 4 1/2 inches, inside dimensions, were constructed of double-wall hardwood. The chambers were arranged in three rows of five and each connected in parallel to an olfactometer constructed as described by Pfaffman, Goff, and Bare (5). Some modifications were made to improve thermal control of the air passed to the chambers and to achieve a parallel equal flow of air to each chamber. For the data to be reported, the olfactometer was used only to filter room air and control ventilation (2500 cc per minute) and temperature (75°F) in the feeding chambers.

A straight runway was constructed, painted flat black, and housed in an adjacent sound-dampened room. The runway had a 9 7/8-inch starting chamber, a 48 3/4-inch runway, and a 10 1/2-inch goal box. The runway between the starting and goal boxes could be inclined to provide runway slopes from the horizontal to about 75°. Hardware cloth tacked to the floor permitted easy traction for the animal. Hinges permitted tilting the running area while keeping the starting and goal boxes horizontal. The starting and goal box doors were guillotine-type, operated manually but remotely with strings. The top of the runway was covered with lucite.

A microswitch and a photoelectric, latching, relay circuit controlled a .01-second timer which started when the starting door opened and stopped when the animal broke a light beam one inch past the starting door. This provided a measure of starting latency. A second photocell started a second timer when the animal reached a point in the runway 18.5 inches past the starting door. This timer was stopped when the animal broke a third beam 20 inches past the second one. Thus, running was measured for a 20 inch-segment located in the center of the runway.

Purina Mash was used for food. This was prepared as a damp mash and served in a glass dish. Trinitrobenzine aniline complex (TNBAC) was used as the repellent. This was dissolved in methanol, then mixed with a blender into the mash. Uncontaminated food was prepared the same way, using the methanol alone. Both preparations were made 24 hours before use to permit the methanol to evaporate before serving.

2. *Procedures*

Eighteen albino rats, 90 days old at the start of the experiment, were placed on a 23 1/2-hour food deprivation schedule in their home cages. During this time, and for the remainder of the experiment, they were permitted 30 minutes within which to eat. Water was available *ad lib.* Body weights, measured to one gram, were always obtained just before feeding.

On Days 8, 9, and 10 the animals were put in the runway one at a time for 5 minutes before their daily ration. Both doors were open. Two small, dry, 45 mg food pellets were placed in the goal box. The animal was permitted to explore without restriction.

On Day 11 the animals were put in the starting box with the door closed and given five successive runs to a reward of five 45 mg pellets each time. The runway was at zero degrees inclination. Following this they were put into the feeding chambers for their 30-minute eating period. From this point on no food was ever given in the home cage. After 22 days on this schedule, following weighing, the animals were placed *first* in the feeding chambers for 25 minutes. After this each animal received one trial in the runway to a dish of damp mash in the goal box. Thus, the 30-minute eating period now consisted of two parts, a 25-minute feeding chamber period followed by a 5-minute eating period in the goal box of the runway. The start of the 25-minute period was staggered for each animal so that there was no delay between the 25-minute and the 5-minute periods except the time required to remove the animal from the feeding chamber, put it in the starting box, and for it to run to the goal. Excluding the runway time, the maximum interval was about 15 seconds.

The last schedule was maintained for 16 days. Then, with the use of the individual means across Days 15 and 16, the animals were matched as closely as possible into groups of equal group-mean body weight, food consumption, and running time. From this point on each group was characterized by a different concentration of TNBAC in the food presented in the feeding chamber. The food in the runway was always uncontaminated.

In order to minimize the time between the feeding and runway procedures, it was possible to handle only 12 animals at a time at the start of the runway training. As a result the matching procedure was applied only to these 12. The remaining six animals were delayed in starting the experiment until approximately one month after the others had been started. They were, however, maintained on the 23 1/2-hour deprivation period over the same time period. Of these animals, one died, leaving five animals for the group.

3. *Experimental Design*

The first group of 12 animals were matched into two groups of six each. One was treated as a 400 ppm TNBAC group, the other as a 250 ppm group. The remaining animals, that started late, were treated as a 100 ppm group.

The experiment proper started with the first day on which matched groups were available. For the 100 ppm group this was the equivalent day counting from their first exposure to the runway. Table 1 describes the experiment from this day on. As shown, Period A was a pretest control period; Period B was a test period during which food in the 25-minute feeding session was contaminated at 100, 250, or 400 ppm according to group. Period C differed only in that the runway was elevated to 45°. Period D was a second uncontaminated period.

TABLE 1
EXPERIMENTAL SERIES

	Days			
Measure	A 1-16	B 17-30	C 31-40	D 41-42
Food	U*	C*	C	U
Runway elevation	25°	25°	45°	45°

* U and C designate uncontaminated and contaminated (TNBAC) food.

C. RESULTS

The basic data were the daily 25-minute and 5-minute food consumptions, body weights, starting, and running times. The 25-minute food consumptions[2] are shown in Figure 1 as means per day of the three groups. The groups may be seen to have been well-matched except for the 100 ppm group whose food consumption was greater during Days 11-15 and less on Day 16 of this period. All groups decreased in food consumption markedly at the start of Period B, the 100 ppm group even more than the 250 ppm group. By Day 19, however, the groups could be ranked clearly by TNBAC concentration in order of increasing food consumption, the 100 ppm group eating the least.

Figure 1 also shows that during Period B all groups tended to recover a portion of their control period food consumption. The 400 ppm group recovered more slowly than did the 250 ppm group and to a lower final level.

---

[2] These food consumptions and those in the goal box were also analyzed with respect to gram/body weight. The same general results were obtained.

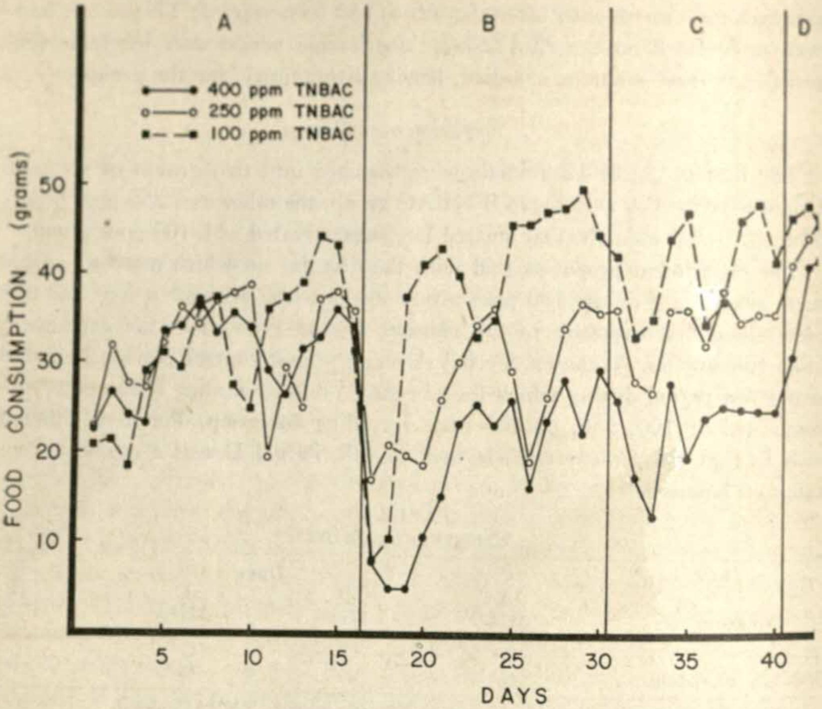


FIGURE 1
MEAN DAILY 25-MINUTE FOOD CONSUMPTION

The 100 ppm and 250 ppm groups recovered completely and, in fact, the 100 ppm group increased its food consumption slightly over its Period A level. All three groups showed a reduction of consumption during the early part of Period C as compared to the end of Period B.

Both the 250 ppm and the 400 ppm groups exhibited a marked increase in food consumption during the uncontaminated Period D; however, the food consumption of the 250 ppm group remained the larger of the two during this period. The figure also suggests the possibility of a slight increase for the 100 ppm group, but this is less clear since there is a tendency for the previous general trend to be rising and since Days 39 and 40 essentially match Period D for this group. Comparison of Period D with the last two days of Period A shows great similarity except for the 400 ppm group on the first of these days.

Figure 2 shows the body weights of the three groups over the course of the

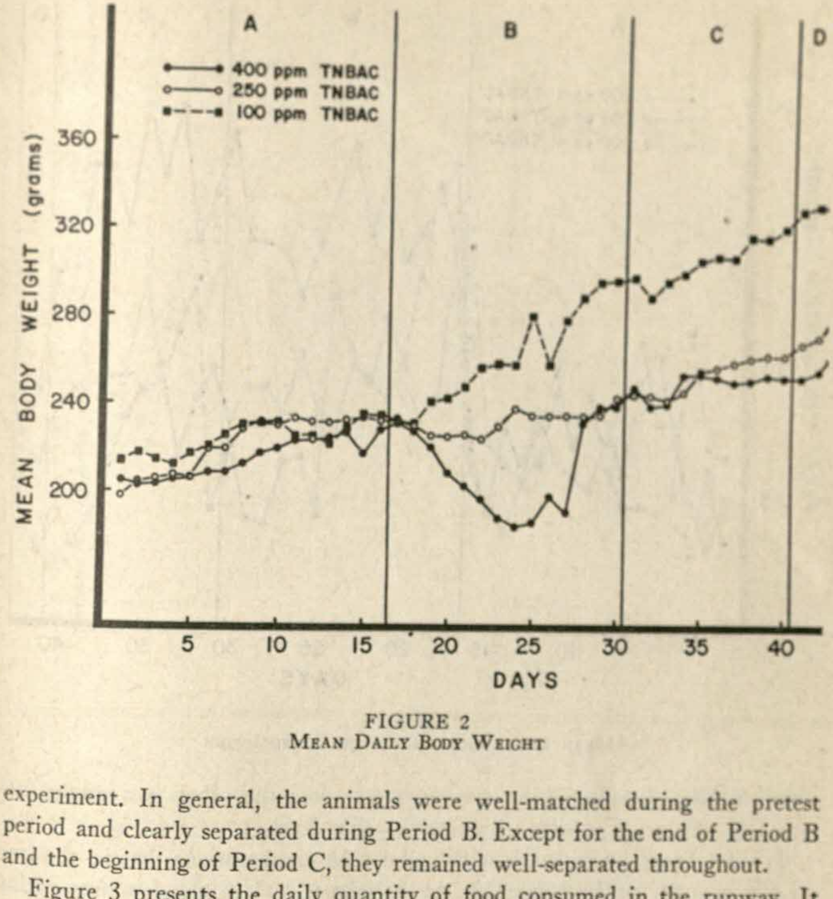


FIGURE 2
MEAN DAILY BODY WEIGHT

experiment. In general, the animals were well-matched during the pretest period and clearly separated during Period B. Except for the end of Period B and the beginning of Period C, they remained well-separated throughout.

Figure 3 presents the daily quantity of food consumed in the runway. It can be seen that the two higher concentration groups were well-matched on Days 15 and 16 whereas the 100 ppm group was not.

During Periods B and C, the 400 ppm group was consistently the greatest consumer of food. From Day 25 and on through Period D, the 250 ppm group consumed more food in the runway, in general, than did the 100 ppm group. Thus, from Day 25 on, except for Days 32 and 33, the three groups are clearly distinguishable, the 400 ppm group being the greatest consumer and the 100 ppm group being the least.

All three groups showed a systematic decrease in food consumption in the runway from about Day 35 to the end of Period C. In Period D, the 400 ppm

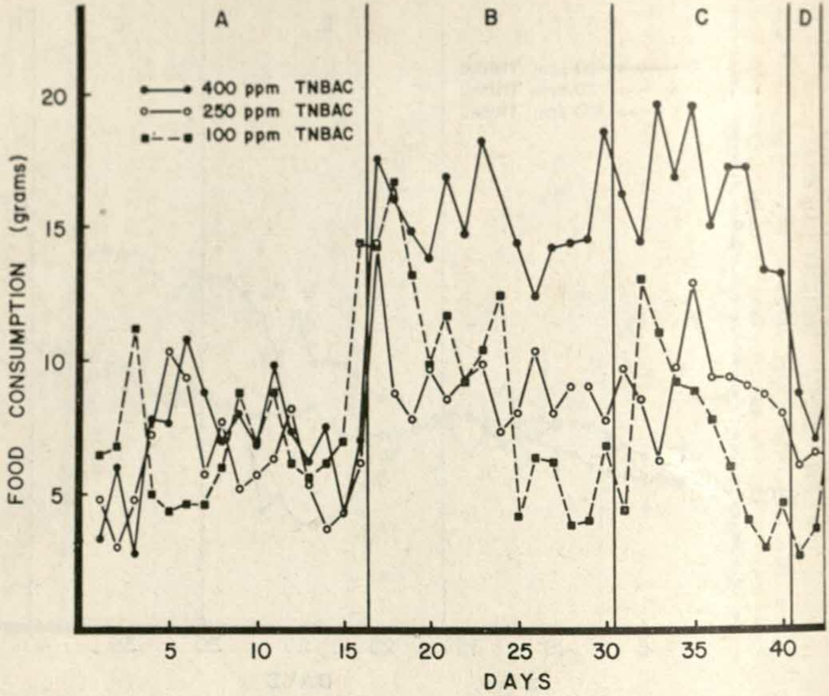


FIGURE 3
MEAN DAILY GOAL BOX FOOD CONSUMPTION

group decreased consumption even more whereas the other two groups apparently reached a possibly temporary, stable level.

It was of interest to consider the total food consumption: that is, the sum of the quantity of food eaten in the 25-minute and in the 5-minute eating periods. These data are shown in Figure 4 in terms of group means. They tend to reflect decreased differences between the groups following Period A as compared to either of the separate food consumptions. Nevertheless the data cannot be interpreted as suggesting a constant total food consumption during any particular series of days since comparison with Figures 1 and 3 suggests relatively stable periods for one or both of the separate measures over the same period of time.

On a few occasions, beginning in Period C, animals in the 100 ppm group did not run up the incline. On such occasions the experimenter arbitrarily removed the animal from the runway when the measure reached 100 seconds.

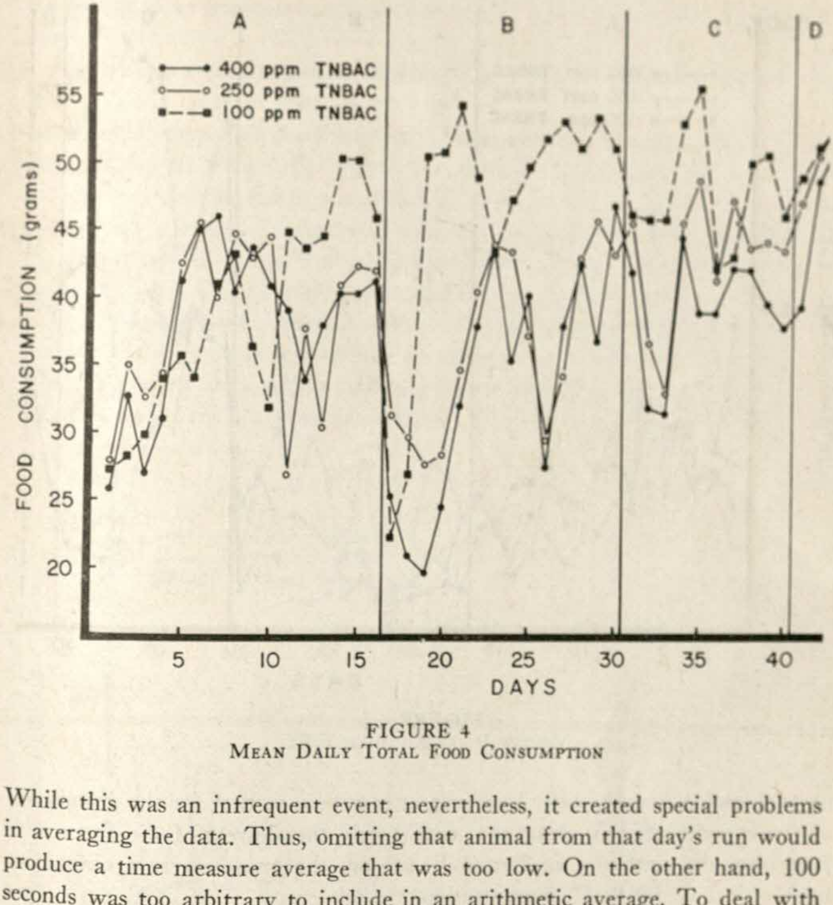


FIGURE 4
Mean Daily Total Food Consumption

While this was an infrequent event, nevertheless, it created special problems in averaging the data. Thus, omitting that animal from that day's run would produce a time measure average that was too low. On the other hand, 100 seconds was too arbitrary to include in an arithmetic average. To deal with this problem, therefore, medians were used as group averages. This statistic has the advantage of allowing the arbitrary limit in the numerical series without being unduly weighted by it.

Figure 5 presents the median running times plotted on a logarithmic ordinate. This type of ordinate was required to include the wide range of values obtained. Inspection of the figure does not suggest systematic differences among the groups during Period A. It can be seen that the 100 ppm group was the fastest at the very end of the period and that this carried over through the first eight days of Period B. On the ninth day of Period B, this group showed a sudden large increase in its running time. The data do not suggest

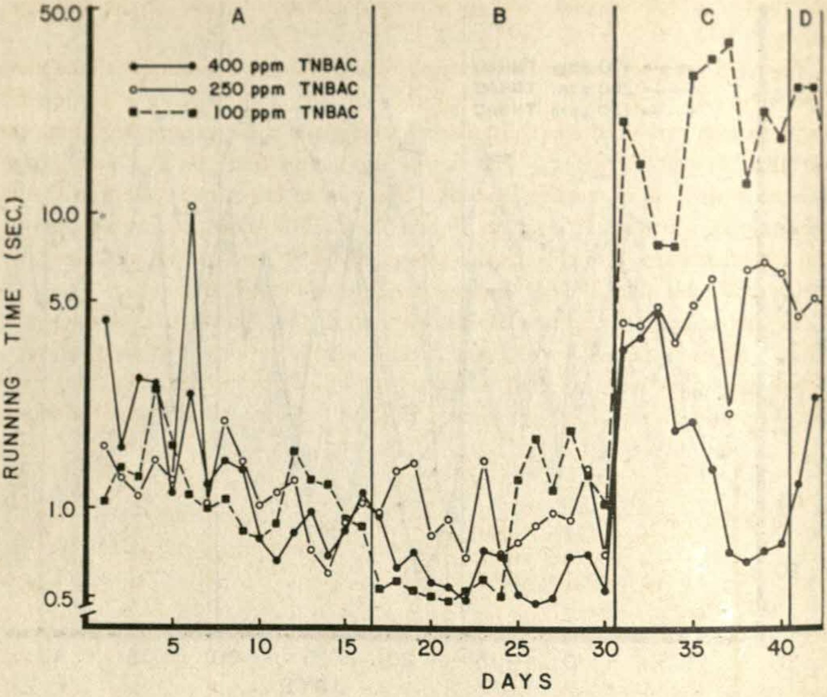


FIGURE 5
Median Daily Running Time

an explanation for this shift. However, it may be seen that for the remainder of the experiment, this group was consistently the slowest of the three and that from the ninth day of Period B and on, the groups were arranged in rank order of TNBAC concentration.

Figure 5 also shows that the general effect of elevating the runway was to produce a very large increase in the differences among the groups and a general slowing down of running. That these effects were not due to inability of the animals to run as quickly as they had in the previous period is seen in the behavior of the 400 ppm group which showed a marked slowing of its running through the first three days of Period B, and then a systematic decrease in running time so that by the end of the period it was running as fast as it had been at the end of the previous period and only .2 second slower than its fastest time in Period A. The other two groups showed no such recovery. The figure also shows that the major effect of removing the contamination from the food was an immediate and increasing reduction of speed of the 400 ppm group.

Figure 6 presents the median starting times, again as a semilogarithmic plot. The figure is clear in showing that the 45° elevation was an experimental factor which produced clear differences among the three groups beginning on the fifth day of Period C. The figure also shows that the 100 ppm group became very slow in starting and that this was an effect not present with the lesser-sloped runway. In fact, in Period A, the 100 ppm group tended to be the fastest starter. Aside from this group, the 250 ppm group was generally a slower starter than the 400 ppm group during Periods B and C.

To investigate the interrelationships among the dependent variables, the average of each animal across Days 39 and 40 was obtained for each measure. These two days were selected as having provided the largest range for each measure, and thereby having provided the best basis for obtaining correlations.

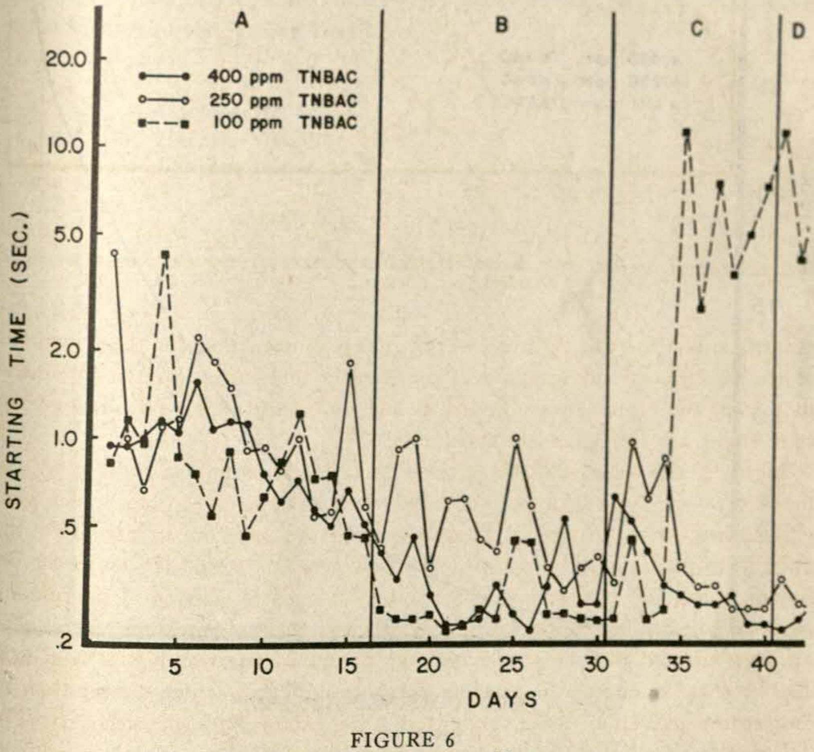


FIGURE 6
Median Daily Starting Time

Figure 7 presents running time as a function of 25-minute food consumption. Each point is the average value of a single animal on Days 39 and 40. The smooth line is a parabolic fit by the method of least squares. The equation which represents the line is

$$t = 1.70 \times 10^{-7.00} \times C^{4.83} \qquad [1]$$

where t is running time in seconds and C is the 25-minute food consumption in grams.

It may be noticed that there are two very poorly fitting points in Figure 7. Otherwise, the obtained values fall reasonably around the fitted line. The index of determination (ratio of the variance of the calculated values to the variance of the observed values) for this fit is .487 which indicates that 49 per cent of the variation in running time is accounted for by the variation in food consumption as represented by Equation 1.

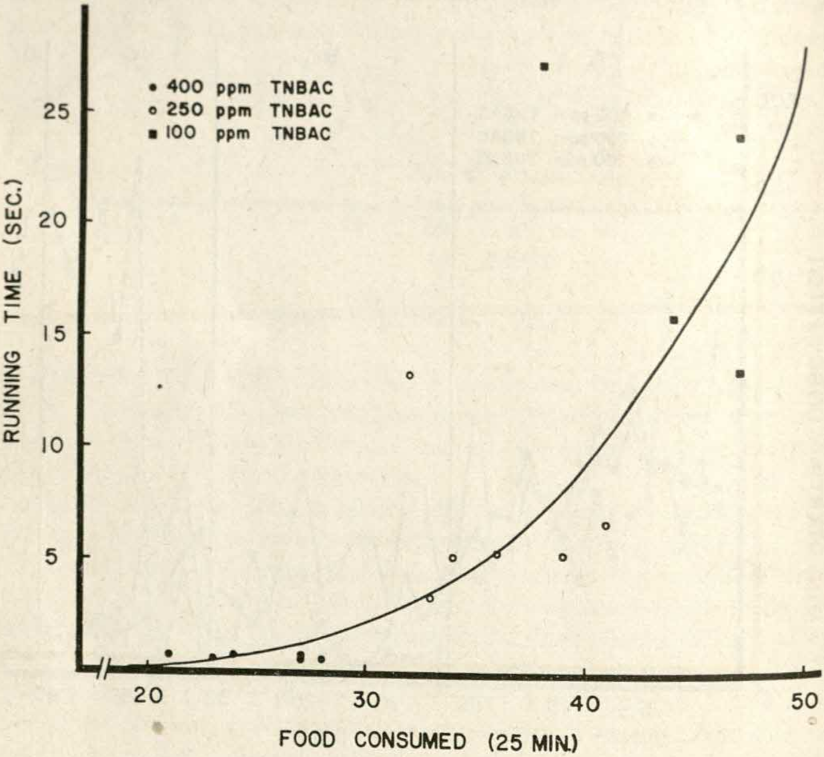


FIGURE 7
RUNNING TIME AS A FUNCTION OF 25-MINUTE FOOD CONSUMPTION

In fitting the points in Figure 7 a number of other equations were tried, all parabolic in nature. Equation 1 was selected as providing the highest index of determination. Yet, it is clear that this expression, logically, must be restricted in generality to the range of food consumptions involved in this experiment, since, as Figure 7 shows, a running time of zero seconds would be expected following a food consumption of approximately 20 grams. A more theoretically reasonable expression would be one which became asymptotic in running time at some value greater than zero seconds. However, for the present purpose reliability of prediction appears more important especially since the asymptote must necessarily depend on the length of runway used.

Figure 8 shows the relationship between food consumption in the goal box (5-minute food consumption) and the 25-minute food consumption. As may be seen, the relationship is negative and linear; the product-moment correlation is .87. Thus, at least within the range studied, the data indicate that the amount of food eaten in the runway can be predicted fairly well from the amount eaten in the feeding chamber.

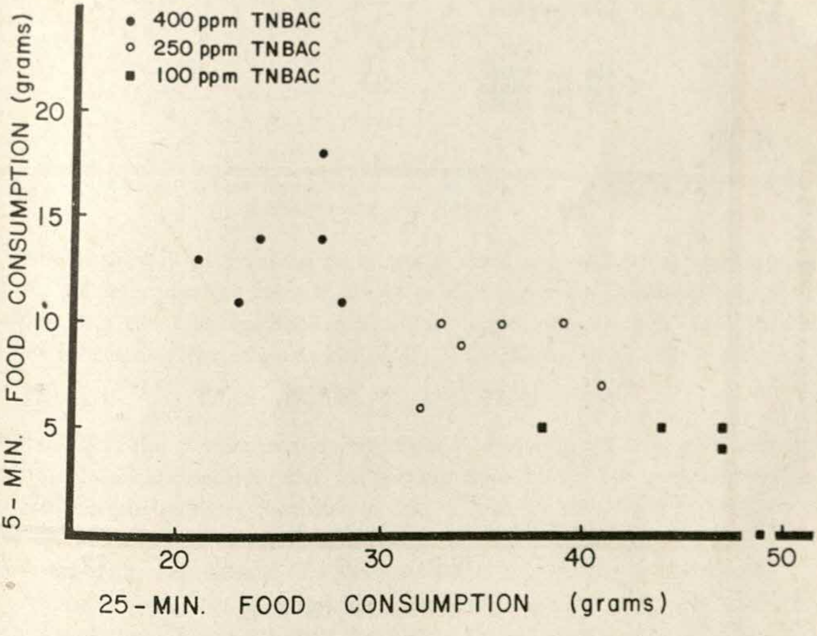


FIGURE 8
Goal Box Food Consumption as a Function of 25-Minute Food Consumption

Figure 9 shows the 5-minute food consumption as a function of running time. The relationship might be expressed as an hyperbola. The single point shown as zero food consumption is arbitrarily placed as this was an instance of a failure to run within 100 seconds. This value was not used in any of the running time curve-fitting procedures (*cf.* Figures 7 and 10). Figure 9 indicates that food consumption in the runway decreased rapidly at first and then slowly approached zero, or some value slightly above it, as running time increased. Most of the variation in food consumption occurred within the running time range of about .5 to 10.0 seconds.

Figures 8 and 9 show that the 5-minute food consumption was related to both the 25-minute food consumption and to the running time. To describe this joint dependency, a surface was fitted to the three quantities. This is shown in Figure 10. The equation was determined by the method of successive approximations knowing from Figure 8 that the relationship with the 25-minute food consumption should be inversely proportional, and from Figure 9 that an hyperbola might describe the relationship with running time.

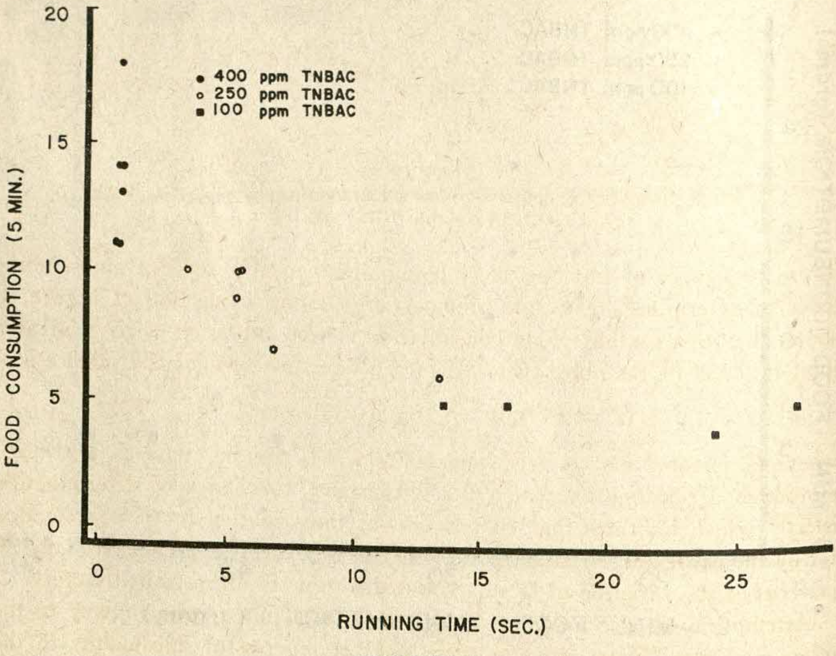


FIGURE 9
Goal Box Food Consumption as a Function of Running Time

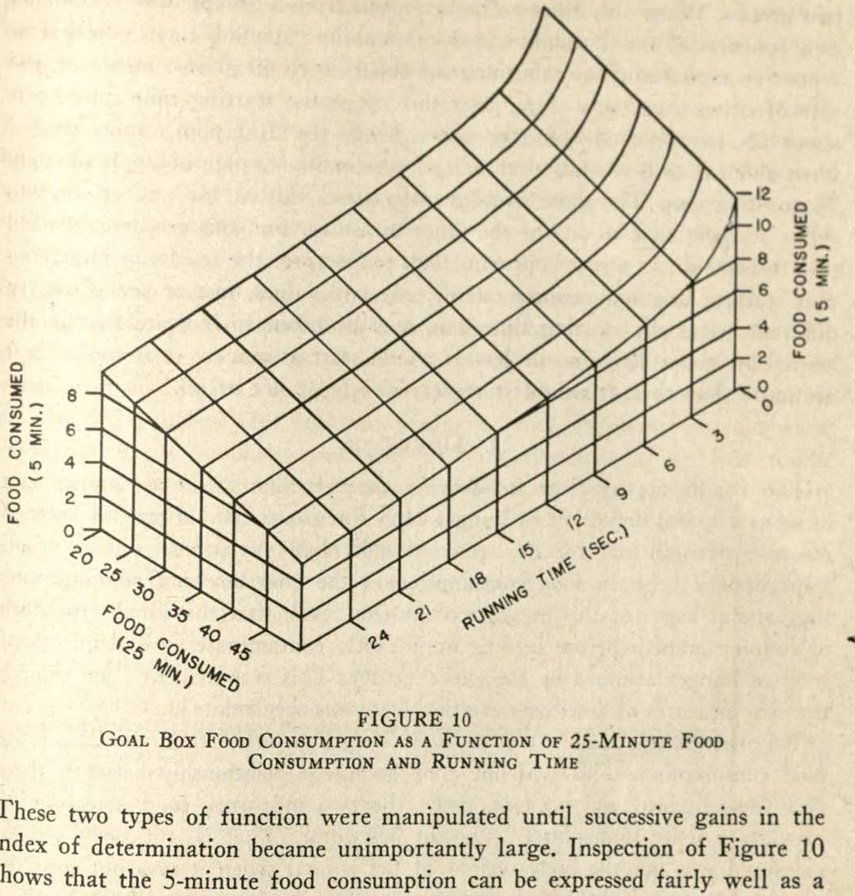


FIGURE 10
GOAL BOX FOOD CONSUMPTION AS A FUNCTION OF 25-MINUTE FOOD CONSUMPTION AND RUNNING TIME

These two types of function were manipulated until successive gains in the index of determination became unimportantly large. Inspection of Figure 10 shows that the 5-minute food consumption can be expressed fairly well as a joint function of the other two variables. The expression for the surface is

$$C' = (17.30t^{-.11}) - (0.14C) - 0.92 \qquad [2]$$

where C' is the 5-minute food consumption, t is running time, C is the 25-minute food consumption, and .92 is a constant. The index of determination is .707 which indicates that 80 per cent of the variation in C' is accounted for by the expressed variation of the other two quantities. This leaves only 20 per cent of the variation of C' unaccounted for due to uncontrolled factors.

Attempts were also made to relate the starting time in the runway to the other measures. These attempts were not very successful due mainly to the very large difference in starting time between the 100 ppm group and the other two groups. When only the two higher concentration groups were considered as a function of the 25-minute food consumption, starting time increased as a positive exponential over the range of about 20 to 40 grams. However, the rate of change was very slow; over this range the starting time range was about .2- to .3-second. On the other hand, the 100 ppm group ranged from about 3 to 8 seconds over a food consumption range of approximately 38 to 48 grams. The same problem, of course, existed for any attempt to relate starting time to one of the other measures. For this reason it seemed most reasonable, as a first approximation, to interpret the results as indicating that starting time may assume one of only two values, fast or very slow. In different terms the starting time data may be taken to indicate that if the animal ate about 40 grams or less, it would start as quickly as it could; if it ate more than that, it would start very slowly or not at all.

D. Discussion

The results appear clear in showing that repellent food acceptance can serve as a useful definition of hunger both for group comparisons of learned response strength and for the study of individual variations. The relationship obtained between food consumption in the goal box and running time suggests, at least for this measure of hunger level, that the usual procedure of running animals before feeding importantly contaminates the definition of level of hunger assumed by the experimenter. This was evident even though the two measures of food consumption were strongly related.

No attempt was made to relate body weight to other than the 25-minute food consumption. This was not done because a relationship between these two was obtained and because, given the two measures, food consumption provides a more immediately relevant descriptor. That is, the body weight measured was not the body weight of the animal when it was put into the runway. Body weight before eating may describe a more general level of hunger as compared to food consumption (or, perhaps, body weight after eating) which describes a more immediate level. On this basis, food acceptance should be more highly related to later performance. A second reason for ignoring possible body weight relationships was that the instrument used to measure body weight provided a relatively gross measure compared to the other measures taken. The possibility of obtaining functional relationships, therefore, was considerably less.

The results suggest that to relate learned response strength to hunger level requires consideration of the amount of work required to make the response. A comparison of different drive levels could fail to show an effect on response strength either because too little or too much work was required. The interaction of drive and work has been an interest of behavior theory for a long time [see Hull (4)]. The present results suggest an approach to the study of this interaction through factorial, manipulation of hunger level, as defined, and runway inclination. This experiment suggests that such studies will show that physical work is not a simple physical variable. The recovery of running speed following the increased runway inclination suggests that a motivational phenomenon is really involved.

In using an acceptance-repellent definition of hunger, it is necessary to avoid chronic toxic effects. A variety of chemical repellents are available which are not toxic at low concentrations and which would appear to provide a suitable range of food acceptance levels. Relatively little really relevant data are available and these only for rodents. There is a need, therefore, for studies of repellency, especially in combination with the effects of deprivation time, in order to establish standard operational procedures for use with different species. In a different context such studies would also appear to have direct relevance to both basic and applied questions of repellency and toxicity. Along with studies of learned response strength they might provide meaningful, sensitive measures to relate to the phenomena of interest to physiologists and biochemists concerned with these questions.

REFERENCES

1. BOLLES, R. C. The readiness to eat and drink: The effect of deprivation conditions. *J. Comp. & Physiol. Psychol.,* 1962, **55**, 230-234.
2. ————. Readiness to eat: Effects of age, sex and weight loss. *J. Comp. & Physiol. Psychol.,* 1965, **60**, 88-92.
3. DAVENPORT, D. G., & GOULET, L. R. Motivational artifacts in standard food-deprivation schedules. *J. Comp. & Physiol. Psychol.,* 1964, **57**, 237-240.
4. HULL, C. L. Principles of Behavior. New York: Appleton-Century-Crofts, 1943.
5. PFAFFMAN, C., GOFF, W. R., & BARE, J. K. An olfactometer for the rat. *Science,* 1958, **128**, 1007-1008.
6. TEICHNER, W. H., WAGNER, M. H., & ROWNTREE, T. Recovery of food consumption of rats on forced repellent diets. *J. of Psychol.,* 1966, **62**, 23-31.

Office of Research Coordinator
Tufts University
Medford 55, Massachusetts

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 197-199.

ASSESSMENT OF PHANTOM LIMB SENSATION AMONG PATIENTS WITH LOWER EXTREMITY AMPUTATION*

*Department of Rehabilitation Medicine, State University of New York,
Downstate Medical Center; and Kings County Hospital Center,
Brooklyn, New York*

SAMUEL HIRSCHENFANG AND JOSEPH G. BENTON

A. PROBLEM

In the rehabilitation of the amputee patient, much emphasis has been placed on the sensations from the stump, which give the illusion that the extremity is still intact. This illusory extremity is referred to as the "phantom limb" (1, 4). Following an amputation, patients often refer to "itching," "tingling," or "burning sensation" in a phantom limb. This phantom limb sensation is felt initially outside of the stump, then decreases in intensity, and, before it disappears, is sometimes felt inside the stump. Individuals evidence marked variation in the length of time during which they are aware of the phantom limb, before there is complete absence of the sensation. Often there is conscious and unconscious repression of the fact of amputation, and the patient clings to the picture of what the limb previously was. Body image, with its strengths and weaknesses, and fantasies, will have a definite effect upon this process. There are wide distances and latitudes involved in this area, some of which will be investigated in this study.

B. METHOD

Our subjects were 30 patients with lower extremity amputation, treated in the Amputee Service of the Department of Physical Medicine and Rehabilitation, Kings County Hospital Center. The group consisted of 21 males and nine females between the ages of 16 and 79 years, with a median age of 56 years for males and 66 years for females. The etiological factors related to the amputation were (*a*) atherosclerosis, (*b*) diabetes mellitus, and (*c*) trauma. Patients with a psychiatric history, alcoholism, or organic brain pathology were not included in the sample. Figure Drawings (2) and the Rotter Incomplete Sentence Test—Adult Scale (3) were administered in that order by the same psychologist to each patient upon admission to the Depart-

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 13, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

ment of Physical Medicine and Rehabilitation. The Rotter Incomplete Sentence Tests were scored according to the rules for obtaining a maladjustment score outlined by Rotter and Rafferty. A high score indicates maladjustment.

C. Results

In an analysis of the Figure Drawings, the males drew the same sex first, 81 per cent, as compared to the females, 66 per cent. For both sexes there were no disturbances of body imagery or distortions on the drawings.

An analysis of the Sentence Completion Tests suggested no preoccupation

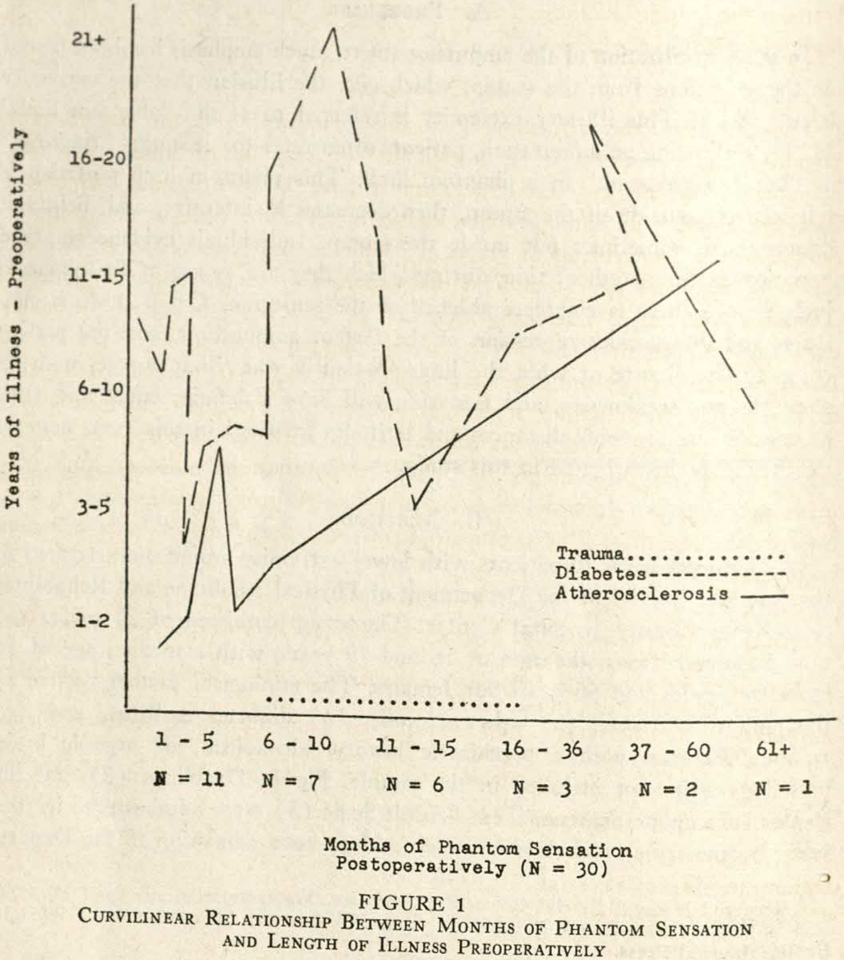


FIGURE 1
CURVILINEAR RELATIONSHIP BETWEEN MONTHS OF PHANTOM SENSATION AND LENGTH OF ILLNESS PREOPERATIVELY

with, or concern over, physical involvement. The mean number of somatic responses given by the males was 4.19 and by the females 4.77, with identical range of responses by both sexes, zero to nine. No evidence of maladjustment was seen in either group. In Figure 1, evaluation was made of the time duration following surgery and the degree and extent of the phantom limb sensation. Those patients who had a traumatic amputation consistently showed the shortest duration of phantom limb phenomenon. Patients who had been ill for the longest period of time, prior to the amputation, evidenced the longest duration of phantom limb sensation. A direct relationship was found between the length of illness, preoperatively, and the duration of the phantom limb sensation postoperatively.

D. SUMMARY

The attitudes, the disturbances in body imagery, and the duration of phantom limb sensation were investigated in a group of lower extremity amputees. Psychological tests, including Figure Drawings and Incomplete Sentence Tests, were administered to patients upon admission to the Amputee Service of the Department of Physical Medicine and Rehabilitation. No significant differences were found on the responses of the Incomplete Sentence Tests for both sexes. The sex drawn first by both sexes indicated the males drew the same sex first (81 per cent), as compared to females (66 per cent). No statistical analysis could be made because of the small number of females in the sample, and further studies are suggested. Traumatic amputees evidenced the shortest duration of phantom limb sensation as compared to the other two groups of patients for whom amputation of a lower limb was made necessary by atherosclerosis or diabetes. A direct relationship is seen between the length of illness, preoperatively, and the phantom limb sensation postoperatively.

REFERENCES

1. FISHER, S., & CLEVELAND, S. E. Body Image and Personality. New York: Van Nostrand, 1958.
2. MACHOVER, K. Personality Projection in the Drawing of the Human Figure. Springfield, Ill.: Charles C Thomas, 1949.
3. ROTTER, J. B., & RAFFERTY, J. E. Manual: The Rotter Incomplete Sentences Blank, College Form. New York: Psychol. Corp., 1950.
4. SIMMEL, M. L. On phantom limbs. *Arch. Neurol. Psychiat.*, 1956, **75**, 637-647.

Department of Rehabilitation Medicine
State University of New York
Downstate Medical Center
450 Clarkson Avenue
Brooklyn 3, New York

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 201-218.

EFFECTS OF STIMULUS CONDITIONS AND RESPONSE CONTINGENCIES ON THE DEVELOPMENT AND MAINTENANCE OF A PERCEPTUAL DISCRIMINATION* [1]

Department of Psychology, University of Missouri at Kansas City

LAWRENCE SIMKINS

A. INTRODUCTION

Several theories have been advanced to account for the development of abnormal behavior within a learning framework [see Dollard and Miller (1), Mednick (5), Mowrer (6), and Skinner (10)]. Although there are some important differences among these various theoretical positions, there are also several areas of agreement. First, all positions seem to imply that the development of deviant behavior is the result of faulty learning experiences. A second implicit assumption is that the development of both normal and abnormal behavior can be most parsimoniously accounted for in terms of learning variables. Third, although there are differences in the conceptualization of these variables, it would appear that a combination of the stimulus or perceptual conditions to which an individual is exposed and the consequences that are made contingent on behaviors emitted under these conditions will determine (*a*) the types of behaviors that are developed, (*b*) the rate of their development, and (*c*) their relative resistance to change. Some of the behavioral characteristics of neurotics include a relatively fixed pattern of response and a failure to adapt appropriately to new environmental conditions.

The persistence of behavior under new environmental conditions may be partly a function of the degree of similarity of stimulus conditions under which the behavior was acquired, and partly a function of the types of response consequences with which the behavior was acquired. Past stimulus conditions and response consequences constitute the behavioral history of an organism. Thus the rapidity with which new perceptual discriminations are learned may

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 14, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported by a grant from the National Association of Mental Health. The author would like to extend his appreciation to Robert McCallum, Jack West, and Lynda Wolfe for their assistance in the collection of the data.

be partly a function of the behavioral history of the organism, and partly a function of the conditions under which the new discriminations are learned. Although there have been numerous studies investigating the effects of response consequences [e.g. Lysak (4), Rigby and Rigby (9), and Smith and Hockberg (11)] on perceptual discrimination, there have been relatively few investigations in which the behavioral history of the organism has been manipulated prior to attempting to produce a perceptual discrimination. In the present study the behavioral history will be referred to as the baseline level of response prior to the introduction of a new set of conditions. These baselines were generated under different types of experimental conditions. The study had three aims: (*a*) to determine the relationship between baseline level of response on the development of a subsequent perceptual discrimination, (*b*) to determine if there was any relationship between the conditions under which the baseline behavior was generated and the conditions under which the subsequent perceptual discrimination was developed, and (*c*) to determine if there was any relationship in the behavioral changes produced during a discrimination phase and the maintenance of these effects after the response contingencies had been withdrawn.

B. METHOD

1. *Apparatus*

The *S* was led into the experimental room and seated at the end of a table that was three and one-half feet long. At the other end of the table in front of the *S* was a vertical-light display panel containing 16 lights. The lights were arranged four to a row. Directly in front of the *S* were two buttons. One button was labeled "projection" and above it was a green light. The other button was labeled "response" and above it was a red light. On the desk was also a digital counter that could register points. During positive reinforcement conditions, a press of the response button would activate this counter.

The control room was located adjacent to the experimental room. A one-way observation mirror enabled *E* to observe *S*. In the control room were located all of the timers and control equipment. A bank of 16 slide switches enabled *E* to control the pattern and number of lights that would be shown on the display board. A series of 16 counters recorded the number of responses *S* made in the presence of each pattern of lights that could vary from one to 16. A Hunter shock stimulator that was wired in series with a pulse former was used as the source of electric shock.

Two Hunter Model 111C decade interval timers were used in this study.

One timer controlled the length of time *S* had to respond. In this study, both timers were set at five seconds.

The sequence of events for the *S* was as follows. When *S*'s green projection light was turned on, he pressed the projection button. This turned on a display of lights on the panel board in front of him. At the same time his green light went out and the red light above his response button came on, and he had five seconds to decide whether or not to respond. At the end of five seconds both the red response light and the panel display lights went out. This procedure constituted one trial. The *S* had to wait until *E* turned on the green projection light for the next trial.

A selector switch allowed *E* to program a variety of response consequences. Thus, it was possible to program the registration of points contingent on a press of the response button, or punishment (shock), or no consequence. In avoidance conditioning, the programming could be arranged so that *S* would receive a shock if he did *not* press the response button at the end of five seconds.

2. *Experimental Design*

There were three phases of this experiment: (*a*) baseline phase, (*b*) perceptual discrimination phase, and (*c*) extinction phase. All *S*s participated in all phases. Each phase lasted five sessions so that there were a total of 15 sessions, and each session consisted of 80 trials.

a. Baseline phase. Within the baseline phase, there were two sets of stimulus conditions and three types of response consequences. The two types of stimulus conditions were (*a*) lights and (*b*) no lights. In the lights condition a certain number of display lights went on each time *S* pressed the projection button. The patterns of lights were randomized as much as possible. In the no-lights condition, the display lights never went on.

The three types of response contingencies were (*a*) positive reinforcement, (*b*) avoidance conditioning, and (*c*) 50 per cent positive reinforcement and 50 per cent punishment (50/50). In the positive reinforcement condition, *S* received a point for all responses. In the avoidance condition, *S* received a shock every time he did not press the response button within five seconds after he pressed the projection button. In the 50/50 condition, *S* received shock for 50 per cent of his responses and points for 50 per cent of his responses.

Thus during the baseline phase there were two types of stimulus conditions and three types of response contingencies. The *S*s were randomly assigned to one of the six possible combinations of stimulus and response conditions. It is also important to note that in the lights stimulus condition during the baseline phase the response contingencies were independent of the number of lights on the display panel.

b. Perceptual discrimination phase. Following the baseline phase, the perceptual discrimination phase was begun. During the discrimination phase, all *S*s were presented a display of lights when the projection button was pressed. In this task, any display pattern consisting of nine or more lights was considered the S^D; eight lights or less was considered the S^{Δ}[2]. Within the perceptual discrimination phase, there were four types of response consequence: (*a*) R + O, (*b*) A + O, (*c*) R + P, and (*d*) A + P. In the R + O condition the *S* received points for each response in the presence of the S^Ds and there were no consequences for a response in the S^{Δ}s. In the A + O condition, *S* received a shock if he did not respond in the presence of the S^Ds but there were no consequences for not responding in the S^{Δ}s. In the R + P condition, *S* received one point for each button response in the presence of the S^Ds and shocked for each response in the S^{Δ}. In the A + P condition, *S* received shock if he did not respond in the presence of S^D and received shock each time he responded in the presence of S^{Δ}.

c. Extinction phase. During this phase, all response contingencies were withdrawn. That is, neither points nor shock were contingent on any of *S*s' responses.

The experiment consisted of a two by three by four factorial design. That is, there were two baseline stimulus conditions (lights and no lights); three baseline contingency conditions (positive reinforcement, avoidance conditioning, and 50/50); and four perceptual discrimination contingency conditions (R + O, A + O, R + P, and A + P).

d. Subjects. The *S*s were drawn from introductory psychology courses at the University of Missouri at Kansas City. A total of 72 *S*s participated in the study. Three *S*s were randomly assigned to each of the 24 cells of the experiment.

C. PROCEDURE

1. *Training*

Part of the first session consisted of a training period in which *S* was acquainted with the procedure of the experiment. The *S* was led into the experimental room and seated at the desk in front of the display panel of lights. The shock electrodes were placed on the index and middle fingers of his left hand. The following instructions were given at this time:

---

[2] An S^D is any stimulus or class of stimuli in the presence of which responses by the *S* are reinforced. An S^{Δ} is any stimulus or class of stimuli in the presence of which responses are not reinforced.

> This is an experiment in visual perception. Before we actually begin it is necessary that we give you some practice so that you know exactly what we want you to do. Here are two lights on your desk and two buttons. One button is labeled "projection" and the other is labeled "response." When the green light is on, push the projection button. When the green light goes off, the red light will come on and will remain on for five seconds. During this time you make a decision as to whether you should press the response button or not press it. If you decide to press the response button, you may do so only when the red light is on. Never press the response button more than once or when the red light is off. Never press the projection button when the green light is on.

Following the instructions, a practice session of at least 25 trials began. It was necessary that *S* clearly understood the instructions that he was only to press the response button once in any trial and that he was only to press the response button when the red light was on. The practice sessions continued until *S* reached a criterion of 10 consecutive trials without committing any of the above errors.

2. *Shock Threshold*

Following the training session, a determination of *S*'s shock threshold was made. This procedure was used on all *S*s regardless of whether or not they were actually to receive shock during the experiment. The *E* turned on the Hunter shock stimulator and slowly increased the variac setting in half-unit steps until the point was reached where the *S* lifted his hand. The dial setting of the Hunter stimulator and the intensity of the current in milliamps was then recorded. If the *S* was to receive shock, the same settings were used in each successive session.

3. *Baseline Sessions*

After the shock threshold was made, instructions were given to the *S* dependent upon the baseline stimulus group to which he was assigned. If he was assigned to the no-lights group, the following instructions were given:

> From now on we will be using the same procedure. Whenever the green projection light comes on, press the projection button. You will then have five seconds to decide whether or not you wish to press the response button. If you decide to press the response button, press it only once in a trial. Also, do not press the response button after the red light goes out. The procedure will be altered somewhat now in that it will be possible for you to get points or shock. Sometimes when you press the response button, a point will register on this counter. At the end of the session, you will receive a penny for each point that is registered. Sometimes when you press the response button, you will receive a shock and

> sometimes when you don't press the button you will receive a shock. I cannot tell you the basis on which you will receive shock or points. The idea, of course, is to try to get as many points as possible and avoid getting shock.

These same instructions were repeated every day for the rest of the experiment. The *S*s that were assigned to the lights group were given the same basic instructions except that the following modifications concerning the presence of the display lights were included.

> Whenever the green projection light comes on, press the button. At this time a display of lights on the board in front of you will come on. . . . I cannot tell you the basis on which you will receive points or shock. Perhaps looking at the light display will help you make your decision. . . .

During the baseline stimulus conditions, if the *S* was assigned to the positive reinforcement condition, he received a point for every button pressed; if the *S* was in the avoidance condition, he received shock every time he did not press the response button; and if he was in the 50/50 condition, 50 per cent of his responses were followed by shock and 50 per cent were followed by points.

4. *Perceptual Discrimination Sessions*

During this phase, *S*s continued to receive the same instructions as they had during the baseline phase. However, the contingencies of their responses may have differed from the baseline contingencies dependent upon the group to which they were assigned. In addition, the contingencies were now all dependent upon the number of lights that were presented in the display panel.

In all conditions, pressing the response button in the presence of the S^D (any combination of nine or more lights) was considered a correct response; pressing in the presence of the S^Δ was considered incorrect. The contingencies of *S*s' responses were dependent upon that one of the four contingency subgroups to which they were assigned (R + O, A + O, R + P, and A + P).

5. *Extinction Phase*

During the extinction phase the same basic instructions were repeated to the *S*s and the light displays continued to be present. There were, however, no consequences programmed for any of the *S*s' responses.

D. RESULTS

1. *Baseline-Response Frequency*

A graph showing the frequency of button-pressing by the various treatment groups during the baseline phase is presented in Figure 1. An analysis of

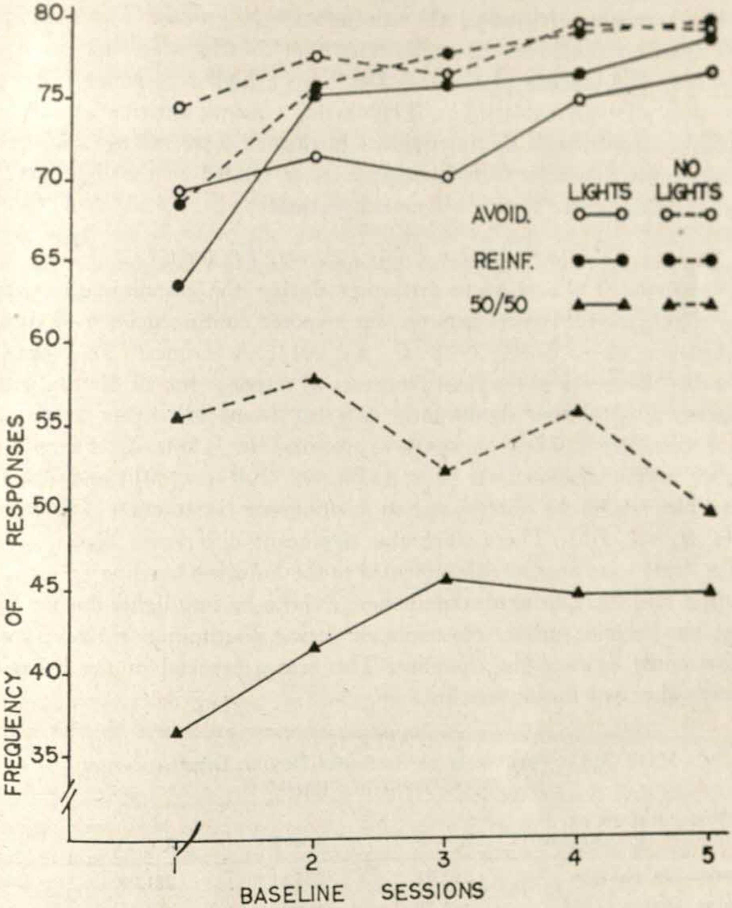


FIGURE 1
Mean Responses Per Session During Baseline Phase

variance for the baseline conditions revealed significant differences (*a*) between baseline stimulus conditions ($F = 10.25$, 1/66 *df*, $p < .01$), (*b*) between baseline response contingencies ($F = 83.34$, 1/66 *df*, $p < .01$), and (*c*) within sessions ($F = 10.74$, 4/264 *df*, $p < .01$). The *S*s in both the avoidance conditioning and the positive reinforcement groups behaved in a similar fashion. After the first session, *S*s in both of these groups emitted close to 100 per cent response. This is in marked contrast to the 50 per cent reinforcement, 50 per cent punishment group (50/50) who typically had a much lower response frequency than the other two groups. The significant baseline stimulus conditions effect indicated that *S*s who were not presented lights during the baseline phase tended to have a higher response frequency than *S*s who were presented lights. This finding was not anticipated and may be partially a function of the instructions in that *S*s were led to believe that the response contingencies were in some manner related to the light pattern display and thus were "testing" various hypotheses.

2. *Discrimination Phase-Response Frequency*

The analysis of the response frequency during the discrimination phase showed significant differences between the response contingencies used during discrimination ($F = 29.68$, 3/48 *df*, $p < .01$). A Duncan analysis (2) showed that the means of the total frequency of response for all discrimination contingency groups were significantly different from each other at the .01 level of confidence. These means are presented in Table 1. There were significant within session effects ($F = 15.58$, 4/192 *df*, $p < .01$) and a significant within session by discrimination contingency interaction ($F = 3.65$, 12/192 *df*, $p < .01$). There were also significant differences dependent on whether lights were or were not presented to the *S* during baseline ($F = 4.95$, 1/48 *df*, $p < .05$). During discrimination, *S*s who had no lights during baseline exhibited significantly fewer responses during discrimination than *S*s who were presented lights during baseline. This was a reversal of the behaviors that were displayed during baseline.

TABLE 1
MEAN TOTAL FREQUENCY OF RESPONSES DURING DISCRIMINATION AND EXTINCTION PHASES*

Discrimination response contingency (group mean)	R + P	A + P	R + O	A + O
Discrimination phase	192.95	225.61	281.00	366.60
Extinction phase	163.50	217.00	215.78	325.94

* In the Duncan test, any two means not underscored by the same line are significantly different from each other at the .01 level of confidence.

Whereas the presence or absence of lights during the baseline phase affected response frequency during discrimination, the baseline response contingencies and the baseline response frequency did not have a significant effect on the discrimination response frequency.

3. *Extinction Phase-Response Frequency*

An analysis of variance of the response frequency during extinction showed significant differences between discrimination contingencies ($F = 12.94$, 3/48 *df*, $p < .01$). A Duncan analysis comparing the means of total response for the four discrimination contingency groups is presented in Table 1. This analysis showed that the R + P group had the lowest level of response and the A + O group the highest frequency of response. The R + O and the A + P groups showed intermediate rates of response and were not significantly different from each other. It is fairly evident that the contingencies programmed during the discrimination phase affected the rate of response decrement. A graph showing the mean frequency of response per session during both the discrimination and extinction phases by the four discrimination contingency groups is presented in Figure 2.

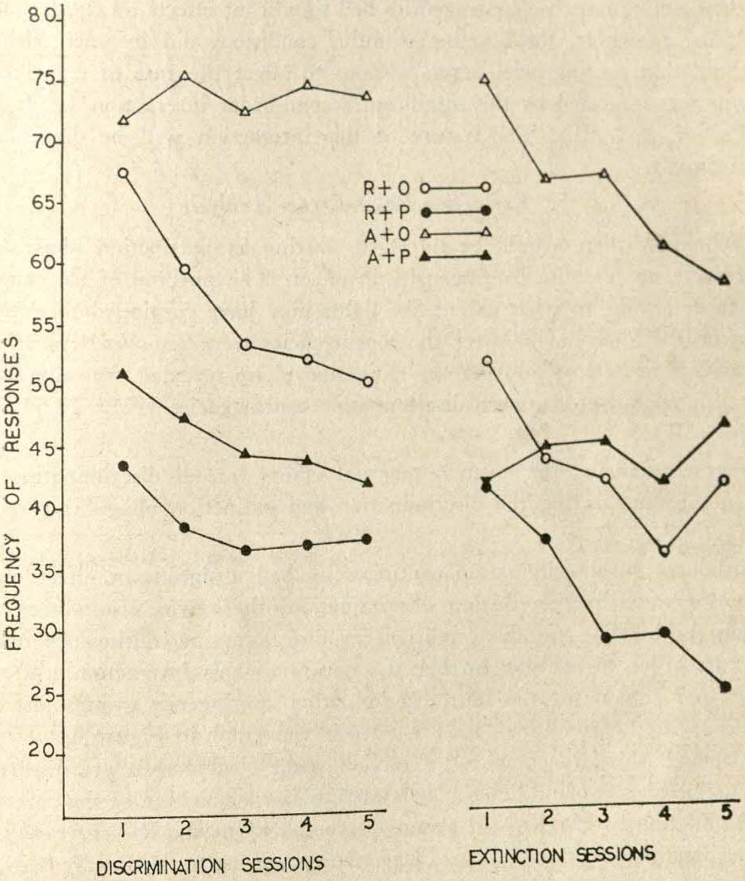


FIGURE 2
MEAN RESPONSES PER SESSION DURING DISCRIMINATION AND EXTINCTION PHASES

4. *Discrimination Phase-Errors Analysis*

An error is defined as either a response in the presence of the S^{Δ} or the lack of a response in the presence of S^{D}. An analysis of variance showed significant differences between discrimination contingency groups in the number of errors ($F = 14.07$, 3/48 *df*, $p < .01$) and in the rate of error reduction across sessions ($F = 89.25$, 4/192 *df*, $p < .01$). A Duncan analysis showed no significant differences between the A + P and the R + P groups. However, the differences between these two groups and the remaining contingency groups were all significant beyond the .01 level of confidence.

In this analysis, neither the baseline stimulus conditions (lights *vs.* no lights) nor the baseline response contingencies had significant effects on the frequency of errors. However, the baseline stimulus conditions did interact with the discrimination contingencies across sessions to affect the rate of error reduction, as was indicated by the significant second-order interaction ($F = 2.30$, 12/192 *df*, $p < .01$). The nature of this interaction will be discussed in detail shortly.

5. *Extinction Phase-Errors Analysis*

Technically there would be no errors during the extinction phase, since there were no response contingencies in effect. The purpose of this analysis was to determine to what extent the lights may have continued to exert an effect on the *S*'s behavior after the contingencies were removed. An analysis of variance of "errors" during the extinction phase revealed the only significant differences were between discrimination contingencies ($F = 24.64$, 3/48 *df*, $p < .01$).

A graph showing the mean number of errors for all discrimination contingency groups during the discrimination and extinction phases is presented in Figure 3.

Although the discrimination contingencies had a significant effect on the rate of error reduction during discrimination, there was also a significant sessions by baseline-stimulus-conditions by discrimination-contingency interaction. In order to analyze further the nature of this interaction, the mean errors per session for the four discrimination contingency groups and their baseline stimulus conditions are presented in the graph in Figure 4.

Although this is a complex interaction, some clear trends are discernible. A baseline history of no lights was related to fast acquisition of the perceptual discrimination by the R + O groups as compared to the R + O group that had a baseline history of lights. These two groups are more divergent in their performance in the initial sessions than in later sessions, suggesting perhaps a

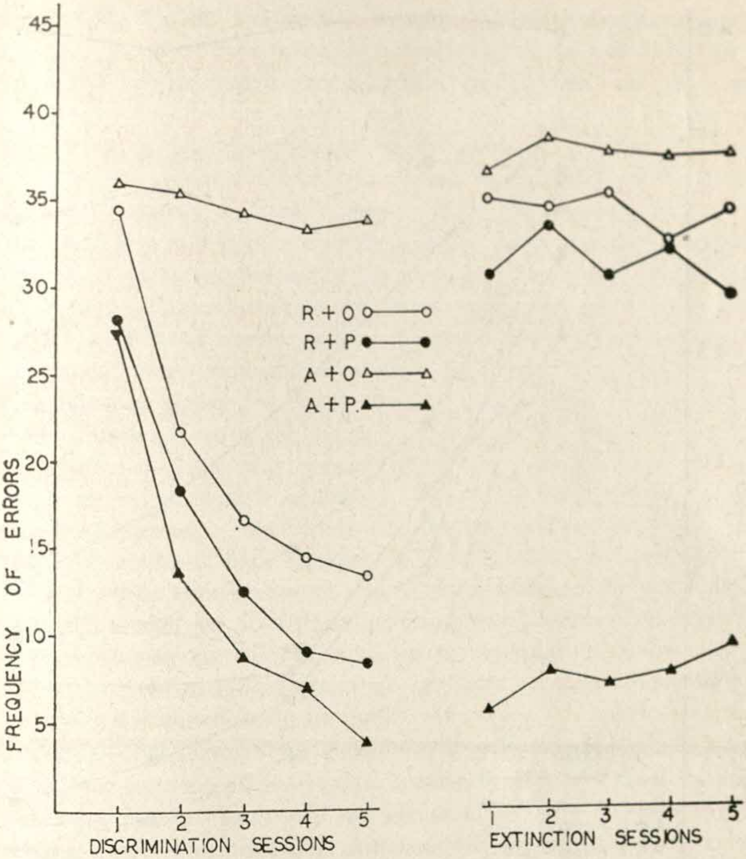


FIGURE 3
MEAN ERRORS PER SESSION DURING DISCRIMINATION AND EXTINCTION PHASES

temporal limitation of the baseline stimulus conditions effect. This same finding is characteristic of the A + P groups in which a history of no lights was related to faster learning than a history of lights. On the other hand, the reverse finding is true of the R + P groups in that those *S*s with a baseline history of lights learned more rapidly than *S*s with a baseline history of no lights. It is somewhat difficult to account for these findings. Generally speaking, it might be anticipated that a baseline history of lights may impede subsequent discrimination, since during the baseline the *S* has learned "erroneous" responses. The performances of the R + O and the A + P groups would tend to confirm this hypothesis, whereas the performance of the R + P

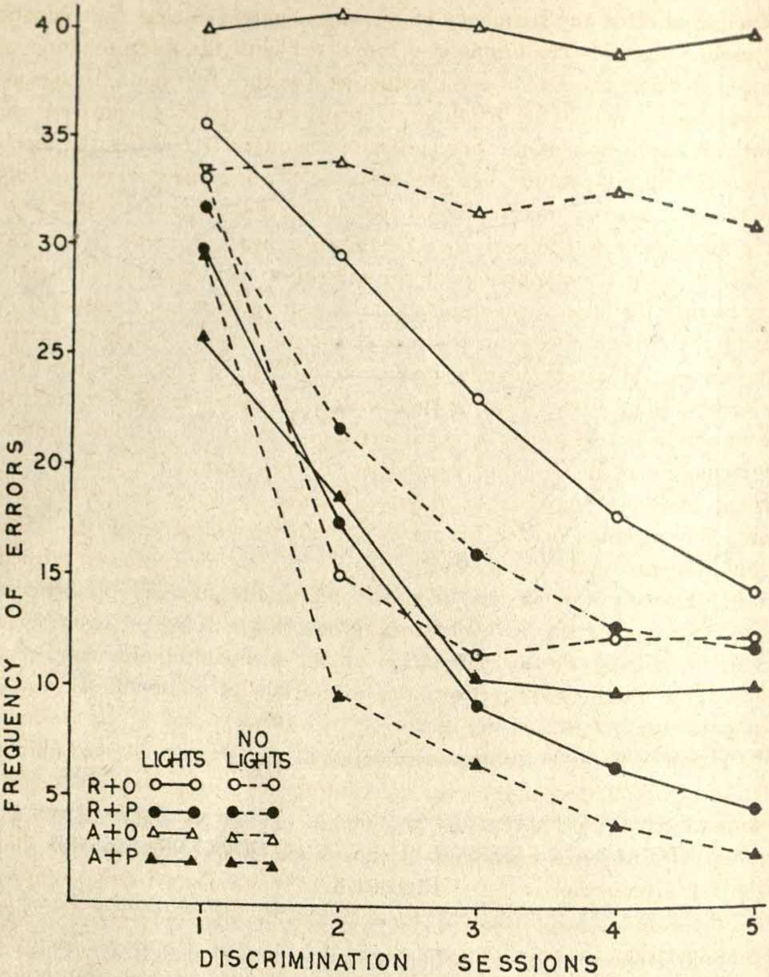


FIGURE 4
Mean Errors Per Session During Discrimination Phase

group would be in contradiction to such a hypothesis. In summary, then, it would appear that the baseline history of stimulus conditions does both affect and interact with discrimination contingencies, but that the nature and reasons for this interaction are not clearly understood.

E. Discussion

The analyses of the data indicate that the presence or absence of lights in the baseline affected the frequency of responses emitted during discrimination. The baseline stimulus conditions also interacted with the discrimination contingencies to affect the rate of error reduction. On the other hand, the response contingencies by which the baseline behavior was established did not affect the rate of acquisition of the perceptual discrimination. Neither the baseline stimulus conditions nor the baseline response contingencies were related to the response frequency or error frequency during the extinction phase.

In general it would appear that the baseline history of the *S* had but a minimal effect on the acquisition of the subsequent perceptual discrimination. It may be that the procedures employed in this study were not sensitive enough to detect the effects of baseline history. A second possibility may be in terms of the amount of discriminable changes in procedures from the baseline to the discrimination phase. Two of the baseline groups were on essentially 100 per cent schedules. During the discrimination phase the potential number of reinforcements available was reduced by 50 per cent (S^Ds and S^Δs were presented an equal number of times). Viewed in these terms, *S*s whose response consequences in baseline were 50 per cent reinforcement and 50 per cent punishment should have acquired the discrimination at slower rates than the other baseline contingency groups. This should have been particularly true in subgroups of the 50/50 baseline group that had R + P contingencies programmed during the discrimination phase. Although there were general trends in the predicted direction, the magnitude of these effects was not sufficient to reach statistical significance.

The major differences in the acquisition of the perceptual discrimination appear to be a function of the response contingencies programmed during the discrimination phase. The R + P and the A + P groups had approximately the same rates of error reduction. Although the R + O group also showed a progressive reduction in the frequency of responses and in the number of errors, the rate of acquisition of the perceptual discrimination was significantly slower than that of either the R + P or the A + P groups. There were many *S*s in the R + O group who failed to acquire perfect discrimination within five sessions and there were a few *S*s who maintained a 100 per cent response frequency. It is speculated that these *S*s behaved as if they were on a 50 per cent reinforcement schedule and could maximize payoffs by responding every time. The *S*s in the A + O group never acquired the perceptual discrimination. Typically, these *S*s maintained 100 per cent response throughout the discrimination phase.

The contingencies programmed during the discrimination phase also affected the performance during the extinction phase. Both the R + P and the R + O groups showed a progressive decrement in the response frequency. The A + O group also showed a tendency to decrease their response rate during extinction. However, the typical performance of *S*s in this group was to maintain the performance established during the discrimination phase. The A + P group also maintained approximately the same frequency of responses during extinction as had been established during discrimination. Furthermore, an analysis of "errors" during extinction would suggest that the performance of the A + P group is clearly controlled by the lights, since their frequency of errors does not deviate significantly from their discrimination performance.

The results of this study can be partially accounted for in terms of stimulus generalization and discrimination theory. The relationship of these principles to this study can be conceptualized in terms of the degree of similarity throughout the various phases of the experiment. Thus it might be anticipated that changes in performance from baseline to discrimination may be a function of the common stimulus elements. According to the principle of stimulus generalization, behavior learned in the presence of the lights during baseline would generalize to the discrimination task and thus retard the rate of acquisition of the discrimination. The results do not completely support this position. Although the presence or absence of lights in the baseline did affect the response frequency during discrimination, this variable did not have a significant main effect on error reduction. On the other hand, there was a significant albeit somewhat uninterpretable interaction of the baseline lights with discrimination contingencies across sessions during the discrimination phase. Apparently the presence of the lights during baseline had an effect on initial sessions of discrimination (see Figure 4). Again, however, the nature of this interaction would not be in complete agreement with a stimulus generalization interpretation.

The performance during discrimination and extinction would, to some extent, be consistent with discrimination theory assuming that the response contingencies are viewed as response-produced stimuli having discriminative properties as well as in terms of their motivational characteristics. This view can easily be justified by past research. Thus, Muenzinger (8) found faster learning in a visual discrimination task if *S*s were punished for correct responses than if they were not punished. Recently, using an operant conditioning task, Holz and Azrin (3) demonstrated that after correlating electric shock with food reinforcement shock alone produced an increase in the rate of response of pigeons. On the other hand, when a low level of shock was selectively paired with nonreinforcement the shock produced larger rates of response decrement than straight extinction. Thus, electric shock that is traditionally viewed as having motivational or suppressive characteristics also has discriminative properties.

In this experiment, the electric shock and the points probably had motivational and discriminative properties. If the response contingencies in this experiment can be viewed as discriminative stimuli, it would then follow that the rate of acquisition may be a function of the number of discriminative stimuli that are available. For all groups the display panel of lights was one set of "constant" stimuli during both the discrimination and extinction phases. For the R + P and the A + P groups, response stimuli were associated with both the S^Ds and S^Δs. These two groups had significantly faster rates of discrimination than the remaining two groups. For the R + O and the A + O groups, response stimuli were only associated with the S^Ds. The R + O group showed slow acquisition of the discrimination; the A + O group showed no discrimination. On the basis of discrimination theory both of these groups should have been equivalent in their acquisition performance. One reason for the discrepancy might be that in avoidance conditioning the stimulus is contingent on *not* responding whereas in the R + O condition the stimulus is response-produced. A second possibility is that the aversive properties of the shock are such as to increase the probability of response so that *S* does not come into contact with the lack of shock in the S^Δ. In this respect the motivational properties of the shock could also be plausible in accounting for the A + P's performance. The avoidance component would increase the probability of response and the *S* would come into contact with the S^Δ contingency. At the same time punishment in the S^Δ would suppress rate of response and bring him into contact with the S^D contingency.

According to the discrimination theory as advanced by Mowrer and Jones (7), resistance to extinction is a function of the similarity of the acquisition stimuli to extinction stimuli. The more similar the stimulus conditions are in the two situations, the greater the resistance to extinction. The results of this study are consistent within this framework. Neither the A + O nor the A + P groups demonstrated extinction effects. By the end of the discrimination phase both of these groups had learned to avoid shock. Therefore, their responses were no longer associated with any stimulus cues except the display panel of lights. The extinction phase was not associated with any additional discriminative cues that were not present during the final session of the discrimination phase. On the other hand both the R + P and the R + O groups showed extinction effects. During the terminal sessions of the discrimination phase the R + P group were successfully avoiding shock but their responses in the S^Ds still produced points. The withdrawal of the point contingencies during extinction produced an immediate discriminable change and thus may be possibly responsible for the decrement in responding. However, once again the decline in responses could have been equally attributable to the loss of the motivational properties of the points as well as the discriminative cues provided by their withdrawal.

The author has empirically demonstrated differences between four combinations of reinforcement contingencies in the development and maintenance of a perceptual discrimination. Discrimination theory can account for only some of these results. One major difficulty is the inability to specify the limiting conditions for the effects that were produced by the various response contingency combinations. In this respect it is difficult to separate the motivational from the discriminative properties of stimuli. It is quite possible that the rate with which the perceptual discrimination is acquired depends not only on the particular combination of response contingencies but also by the relative "values" of the stimuli that are involved. By manipulating the intensity of shock or the amount of money exchangeable for points it would be possible to generate an entire family of acquisition curves for any given response contingency combination. Thus, it may be necessary to obtain additional empirical data that will establish the parametric functions of the various stimulus combinations used in this study in order to resolve the theoretical issues that have been considered.

In respect to the development of abnormal behavior, several learning theorists [e.g., Mowrer (6) and Mednick (5)] attach considerable importance to the role of aversive stimuli in the development and maintenance of pathological behavior. In this study, the behavior of the *S*s who received electric shock manifested behaviors that are somewhat analogous to neurotic behavior. The *S*s who received the A + O discrimination contingency never acquired the discrimination and generally maintained a high rate of response when the shock contingencies were withdrawn. It is also of interest to note that whereas both the R + P and the A + P groups acquired the discrimination fairly rapidly and showed behavioral similarities during the discrimination phase, they behaved quite differently during extinction. Thus, it is quite possible that although different consequences may produce similar effects during the acquisition of a behavior, differential effects of the procedures may be observed only after the contingencies have been removed. Both the A + O and the A + P groups resembled "neurotic" behavior to the extent that during extinction they behaved as if the contingencies were still in effect. One of the puzzling aspects of neurotic behavior is that it is relatively unchanging and nonadaptive. Thus, the behavior may persist long after the original response contingencies are removed. The results in this study, particularly with the A + P group, would suggest that ultimately environmental stimuli in the presence of which avoidance behavior was generated come to control such "neurotic" behavior. If the environmental stimuli are maintaining the behavior, then "large" changes in stimulus conditions should be accompanied by changes in the behavior. A second possibility would involve attaching new consequences to the behavior in the presence of these environmental stimuli or else reinforcing different behaviors in the presence of the stimuli. Which of these procedures would be most effective in producing behavioral change is open to empirical question.

F. SUMMARY

The purpose of this study was to determine the effects of two classes of variables that may be associated with the development and maintenance of a perceptual discrimination. The two classes of variables were stimulus cues and response consequences. There were three phases in this experiment: baseline, discrimination, and extinction. Each subject participated in all three phases. The perceptual discrimination task consisted of presenting random patterns of lights that also varied in number. The *S* had to learn to make a discrimination based on the number of lights presented.

Since it is possible that the behavioral history of a subject can affect his rate of acquisition of a perceptual discrimination, one purpose of this study was to determine what effect a history of presence or absence of lights and certain response consequences would have on the rate of acquisition of a subsequent perceptual discrimination. The behavioral history was referred to as the baseline phase in this study. During the baseline phase the *Ss'* response contingencies were independent of the number of lights presented. The purpose of the second or discrimination phase was to determine the effects of four different combinations of response consequences on the development of the perceptual discrimination. During this phase the response consequences were contingent on the number of lights presented to the subject. During the third or extinction phase the response contingencies were withdrawn and the purpose was to determine what effects the baseline and discrimination conditions had on the maintenance of the perceptual discrimination.

The results indicated that baseline stimulus conditions (lights *vs.* no lights) affected rates of response during discrimination but did not have significant effects on error reduction or on the extinction phase performance. The response consequences programmed during the baseline had no significant effect on either discrimination or extinction phase performance. The response contingencies programmed during the discrimination phase accounted for the major differences in performance during both the discrimination and the extinction phases. Although some of the results were accounted for within a discrimination theory framework, there were several inconsistencies. One of the problems of interpretation of the results is the difficulty in adequately differentiating the motivational and discriminative properties of stimuli. The possible relationships of these results to the development, maintenance, and modification of deviant behavior were also discussed.

REFERENCES

1. DOLLARD, J., & MILLER, N. Personality and Psychotherapy: An Analysis in Terms of Learning, Thinking, and Culture. New York: McGraw, 1950.
2. DUNCAN, F. Multiple range and multiple F tests. *Biometrics,* 1955, **21**, 1-9.
3. HOLZ, W., & AZRIN, N. Discriminative properties of punishment. *J. Exper. Anal. Behav.,* 1961, **4**, 225-232.
4. LYSAK, W. The effects of punishment upon syllable recognition thresholds. *J. Exper. Psychol.,* 1954, **47**, 343-350.
5. MEDNICK, S. A learning theory approach to research in schizophrenia. *Psychol. Bull.,* 1958, **55**, 316-327.
6. MOWRER, O. Learning Theory and Personality Dynamics. New York: Ronald Press, 1950.
7. MOWRER, O., & JONES, H. Habit strength as a function of the pattern of reinforcement. *J. Exper. Psychol.,* 1945, **43**, 293-311.
8. MUENZINGER, K. Motivation in learning: I. Electric shock for correct response in the visual discrimination habit. *J. Comp. Psychol.,* 1934, **17**, 267-277.
9. RIGBY, M., & RIGBY, W. Perceptual thresholds as a function of reinforcement and frequency. *Amer. Psychol.,* 1952, **7**, 321.
10. SKINNER, B. F. Cumulative Record, New York: Appleton-Century-Crofts, 1959. Pp. 202-219.
11. SMITH, D., & HOCHBERG, J. The effect of "punishment" (electric shock) on figure-ground perception. *J. of Psychol.,* 1954, **38**, 83-87.

Department of Psychology
University of Missouri at Kansas City
5100 Rockhill Road
Kansas City, Missouri 64110

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 219-227.

SEX DIFFERENCES IN THE DIFFERENTIAL RECALL OF TABOO AND NEUTRAL WORDS*[1]

American International College and Springfield College

George S. Grosser and Anthony A. Walsh

A. Introduction

Attempts have been made to demonstrate Freudian "repression" by measuring the relative retention of pleasant and unpleasant words. If fewer unpleasant than pleasant words are recalled, repression is said to be defined (8, p. 67). Such "repression of recall" experiments were prevalent in the 1930s and 1940s: *viz.*, Meltzer (2, pp. 673-674), Rosenzweig and Mason (2, pp. 674-675), Rosenzweig (4), and Shaw and Spooner (5).

More recently, the protective function of repression has been emphasized, as in the "perceptual defense" experiments, the first of which was by McGinnies in 1949 (2, p. 714). Sexually taboo words were used by McGinnies as the equivalent of the "unpleasant" words that had been the experimental condition in the earlier research, while neutral words were used as his control. Unfortunately, the interpretation of positive results in experiments on perceptual defense has proved to be clouded, due to possible confounding with such other factors as word frequency, set, and response suppression (2, pp. 712-726).

The present experiment combines the older research technique of testing for recall with the modern contribution of using a set of "taboo" words as the experimental condition and a set of neutral words as the control condition; further, male subjects and female subjects were to be compared, since sexually taboo words seem to be more threatening to females than to males (1, 3). It was hypothesized that neutral words would be better recalled than taboo words, but that differential would be more pronounced for female than for male subjects.

B. Method

1. *Subjects*

Thirty-four male and 33 female subjects were chosen from the American International College summer-school population. No matching procedure was

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 21, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This study was carried out by the junior author under the supervision of the senior author, in June, 1963, in fulfillment of the requirements for a Directed Study at the American International College.

used other than equal representation of the sexes. The mean age of the males was 20.47 with the oldest subject being 35 years and the youngest being 17 years of age. The mean age of the females was 20.03 years with the oldest female reporting 27 and the youngest 17 years of age. The overall mean age of the sample, males and females combined, was 20.25 years of age.

2. *Apparatus*

Twenty 3- by 4-inch glass-enclosed slides were utilized. A tachistoscope was set up to present the stimulus words at a standard one-second flash. A blackboard was used to screen the subject from the experimenter and the words were flashed on a white bulletin board. Ten neutral and 10 sexually taboo words were selected from the list compiled by Thorndike and Lorge (7) (see Table 1).

TABLE 1
WORDS UTILIZED IN ORDER OF PRESENTATION
(Numbers correspond to graph code)

1	complete
2	helpful
3	douche
4	sperm
5	religion
6	penis
7	intercourse
8	graceful
9	research
10	analysis
11	logical
12	vagina
13	masturbation
14	wealthy
15	economic
16	homosexual
17	erection
18	friendly
19	pervert
20	prostitute

3. *Procedure*

Both male and female groups were exposed to the same experimental conditions. Each subject was tested individually. The subjects were seated beside the portable blackboard that screened their view of the slides and projector. They were told what they were required to do: i.e., "You are going to be shown a list of words. Study each word carefully and at the end of the presentation of the words I would like you to repeat as many of the words that you can remember." The words were printed on 3- by 4-inch Radio-Mat slide sheets and presented tachistoscopically for a duration of one second. The full test session took approximately four minutes. All the subjects were tested within a continuous six-hour period so as to prevent any confounding due to the spread of information among the campus populace.

C. RESULTS

Two 2 × 10 chi squares (see Tables 2 and 3) were done: one each on the neutral and the sexually taboo recall frequencies for individual words of males and females. The null hypothesis was accepted in both cases indicating that

TABLE 2

CHI SQUARE* TEST OF MALE-FEMALE DIFFERENCES IN RECALL SCORES ON EACH OF THE 10 NEUTRAL WORDS

	Word code**										
Cell frequencies	10	9	8	14	11	2	18	1	5	15	Total
Male fo***	19	14	16	12	10	23	18	22	18	15	167
Male fe†	17.24	16.34	15.43	13.16	11.35	22.24	18.15	21.33	16.79	14.98	
Female fo	19	22	18	17	15	26	22	25	19	18	201
Female fe	20.76	19.66	18.57	15.84	13.65	26.76	21.85	25.67	20.21	18.02	
Total	38	36	34	29	25	49	40	47	37	33	368

* $X^2_{.05} = 16.92$, $df = 9$. Obtained $X^2 = 1.72$ (not significant).

** See Table 1.

*** fo = frequency observed, each frequency observed in the collection of data.

† fe = frequency expected, that frequency one would expect if every observed frequency equaled every other observed frequency.

TABLE 3

CHI SQUARE* TEST OF MALE-FEMALE DIFFERENCES IN RECALL SCORES ON EACH OF THE 10 SEXUALLY TABOO WORDS

	Word code**										
Cell frequencies	20	4	16	19	3	7	17	12	6	13	Total
Male fo***	17	15	22	25	25	24	25	28	29	23	233
Male fe†	16.11	16.68	23.01	27.61	23.01	27.61	21.86	31.07	25.31	20.71	
Female fo	11	14	18	23	15	24	13	26	15	13	172
Female fe	11.89	12.32	16.99	20.39	16.99	20.39	16.14	22.93	18.69	15.29	
Total	28	29	40	48	40	48	38	54	44	36	405

* $X^2_{.05} = 16.92$, $df = 9$. Obtained $X^2 = 6.34$ (not significant).

** See Table 1.

*** fo = frequency observed, each frequency observed in the collection of data.

† fe = frequency expected, that frequency one would expect if every observed frequency equaled every other observed frequency.

the order of presentation had no linear effect on recall. The frequencies of recall for all the words were about equal.

Figure 1 shows that the females consistently had higher recall scores than the males on nine of the 10 neutral words and the groups were tied on one word (analysis).

Figure 2 shows that the males consistently had higher recall scores than the females on nine of the 10 taboo words and the two groups were tied on one word (intercourse).

A Wilcoxon *T* test, done on all *S*s on differences in recall between the sexually taboo and the neutral words, did not reveal a significant difference. For Wilcoxon *T*s done ***separately*** on male *S*s and female *S*s, the results are different. Males recall significantly more sexually taboo words than neutral words, while females recall significantly more neutral words than sexually taboo words (see Table 4).

TABLE 4
WILCOXON *T* TESTS ON DIFFERENCES IN RECALL BETWEEN WORD TYPES

Description	N	T	M_T	SD_T	z
All *S*s	61	-727.5	945.5	139.22	1.57*
Male *S*s	32	-38.0	264.0	53.48	4.15**
Female *S*s	29	$+109.0$	217.5	46.25	2.35**

* $p = .06$.
** $p < .01$.

A Mann-Whitney *U* test done on the difference between the male and female *S*s in ability to recall *all* the words (taboo and neutral) was found not to be statistically significant. The male and female groups did not differ in ability to remember the words shown them.

For Mann-Whitney *U*s done separately on neutral words and on taboo words, the results are different. On neutral words, the female group outscored the males significantly. The male group, however, recalled significantly more taboo words than the females (see Table 5).

TABLE 5
MANN-WHITNEY *U* TESTS CORRECTED FOR TIES
$(N = 67)$

Description	U	$\frac{n_1 n_2}{2}$	SD_U	z
All words	654.5	561	78.08	1.20*
Neutral words	275.0	561	77.17	3.71**
Sexually taboo words	863.0	561	78.72	3.84**

* $p = .12$.
** $p < .01$.

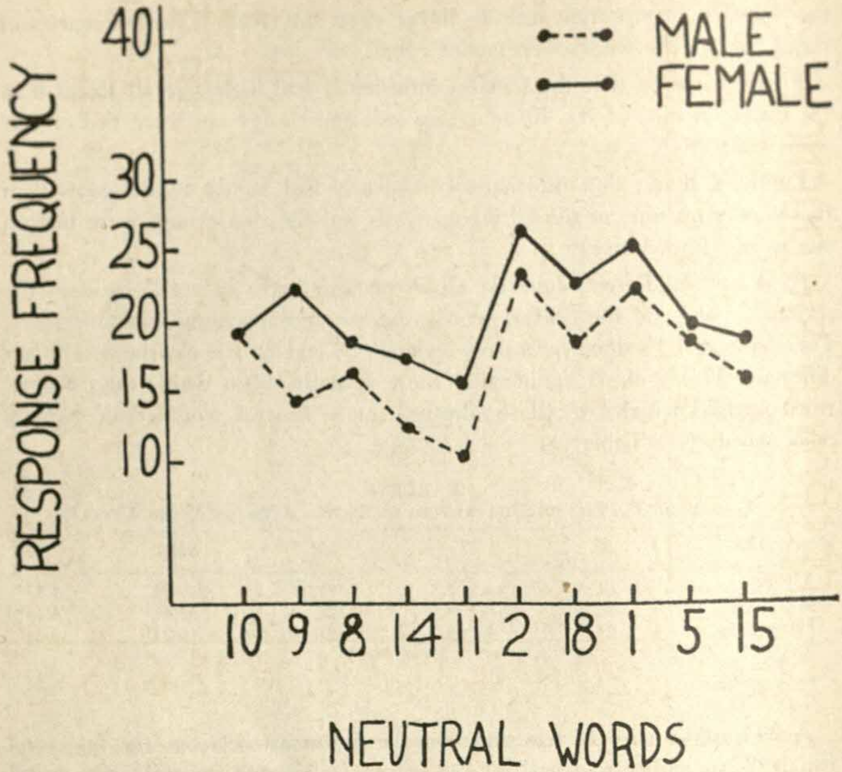


FIGURE 1
MALE AND FEMALE SUBJECTS COMPARED ON RECALL SCORES FOR THE NEUTRAL WORDS
Note that females outscore males on nine of these 10 words. For word code, see Table 1.

D. DISCUSSION

Grosser and Laczek found that, for sexually taboo words, the utterance latencies of 14 male and 10 female *S*s go up in all cases. In the present study, immediate recall tests on 34 males show that the scores are higher for sexually taboo than for neutral words in 28 cases (or 82.34 per cent of the male *S*s). Recall scores for 33 females were found to be lower for sexually taboo than for neutral words in 18 cases (or 54.55 per cent of the female *S*s). Reversals for males (i.e., males scoring higher for neutral words than for sexually taboo words) involved four cases (or only 11.76 per cent of the cases). Reversals for females (i.e., females scoring higher for sexually taboo

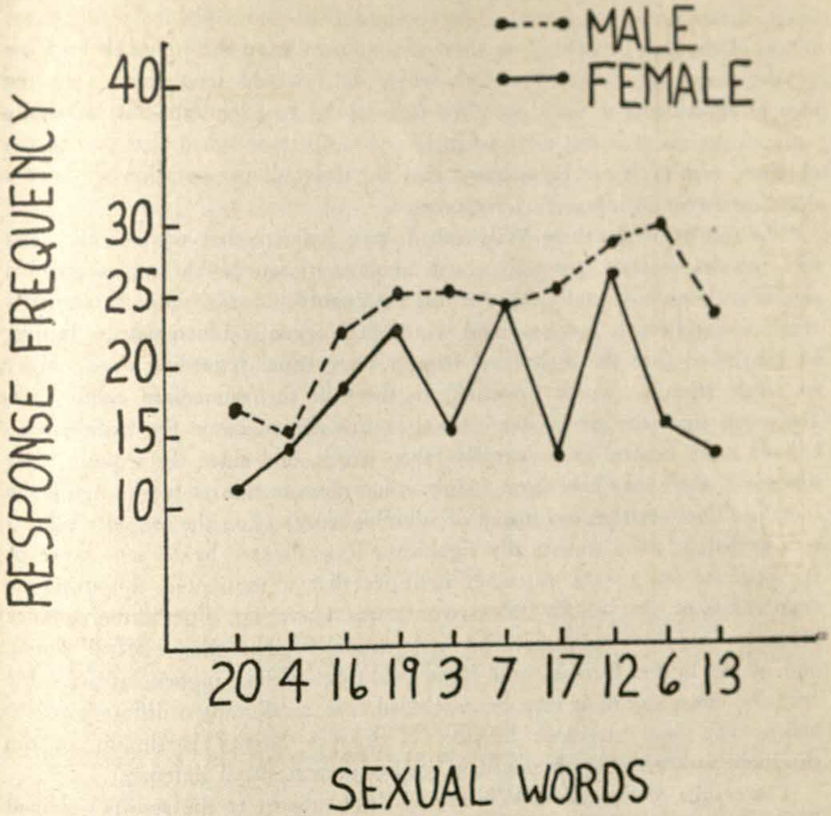


FIGURE 2
Male and Female Subjects Compared on Recall Scores for the Sexually Taboo Words

Note that males outscore females on nine of these 10 words. For word code, see Table 1.

words than for neutral words) involved 11 cases (or 33.33 per cent of the cases). Two males and four females recalled equal amounts of sexually taboo and of neutral words.

The two 2 × 10 chi squares (done on the neutral and sexual words comparing the performance of males and females) were done to ascertain whether or not unidimensionality was reached for the two criterion lists: i.e., whether the sexually taboo words were interrelated on one factor and the neutral words on another factor. While this would be more thoroughly ascertained through a factor analysis than through the chi squares, indications are that two separate dimensions were tapped. The obtained chi squares yielded clear acceptances of the null hypothesis, as there was no significant difference in response by the males and females to each word. This would lend support to the idea of equivalence of word weights. Apparently, two separate but internally equated groups of words were used. It may be further noted that due to the obtained results, it can be assumed that the order of presentation was not a deciding factor in influencing responses.

The results of the three Wilcoxon *T* tests indicate that while both males and females recalled generally equal amounts of words, the males recalled significantly more sexually taboo words than neutral words, while the females recalled significantly more neutral words than sexually taboo words. It may be concluded that the males and females were equal regarding their ability to recall stimulus words presented to them in the immediate recall task. However, since the groups were equal in this ability, since the females verbalized more neutral than sexually taboo words, and since the experimenter was male, there may have been a suppression phenomenon at work among the females. The fact that one group of stimulus words (i.e., the sexually taboo) was verbalized to a statistically significant lesser degree by *S*s who were of the opposite sex to the examiner indicates that a motivated forgetting or suppression of the socially taboo words was operating. Furthermore, since 28 males and 11 females recalled more sexually taboo than neutral words, and all *S*s in the Grosser and Laczek (1) study had higher latencies for sexually taboo words, it may be concluded that recall taps a different mechanism than does utterance latency. (Utterance latency is similar to the threshold measurement employed in studies of perceptual defense.)

The results of the Mann-Whitney *U*s lend support to the results obtained in the Wilcoxon *T*s. It was found initially that males and females did not differ significantly in their ability to recall all the words. Again, it was shown that the two groups of *S*s were of equal ability. It was found further that the neutral words were recalled more often by the females and the sexually taboo words more often by the males.

E. Summary and Conclusions

An experiment was carried out on the recall of sexually taboo words (the experimental condition) and of neutral words (the control condition). Further, male subjects and female subjects were compared. It was hypothesized that neutral words would be better recalled than taboo words, but that this differential would be more pronounced for female than for male subjects. Thirty-four male and 33 female subjects were chosen from the American

International College summer-school population. Twenty 3- by 4-inch glass-enclosed slides were utilized and the words on them were presented tachistoscopically for one second in a setting wherein the subject was screened from a male examiner by a portable blackboard. Ten neutral and 10 sexually taboo words were selected from the list compiled by Thorndike and Lorge (see Table 1). Both male and female groups were exposed to the same experimental conditions and each subject was tested individually. All subjects were tested within a continuous six-hour period to prevent any confounding due to the spread of information among the campus populace.

It was found that while both males and females were equal in recall ability, the females verbalized significantly more neutral words than sexually taboo words, while the males verbalized significantly more sexually taboo words than neutral words. Furthermore, it was found that the order of presentation had no effect on the subject's ability to recall and that the words within each class were from equal populations.

It was concluded that recall taps a different phenomenon than utterance latency does and that this phenomenon may be suppression-like in nature.

REFERENCES

1. GROSSER, G. S., & LACZEK, W. J. Prior parochial *vs.* secondary education and utterance latencies to taboo words. *J. of Psychol.,* 1963, **55**, 263-277.
2. MACKINNON, D. W., & DUKES, W. F. Repression. In L. Postman (Ed.), *Psychology in the Making.* New York: Knopf, 1962.
3. POSTMAN, L., BRONSON, W. C., & GROPPER, G. L. Is there a mechanism of perceptual defense? *J. Abn. & Soc. Psychol.,* 1953, **48**, 215-224.
4. ROSENZWEIG, S. An experimental study of repression with special reference to need-persistive and ego-defensive reactions to frustration. *J. Exper. Psychol.,* 1943, **32**, 64-77.
5. SHAW, F. J., & SPOONER, A. Selective forgetting when the subject is not ego-involved. *J. Exper. Psychol.,* 1945, **35**, 242-247.
6. SIEGEL, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.
7. THORNDIKE, E. L., & LORGE, I. The Teachers Word Book of 30,000 Words. New York: Teachers Coll., Columbia Univ., 1944.
8. UNDERWOOD, B. J. Psychological Research. New York: Appleton-Century-Crofts, 1957.

American International College
Springfield, Massachusetts 01109

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 229-233.

CHILDREN'S BEHAVIOR PROBLEMS AND MOTHER'S AGE*

College of Medicine, University of Iowa

F. Beth Stone and Vinton N. Rowley

A. Introduction

There is currently a resurgence of interest in the effect of family-structure variables—such as birth order, sex of siblings, family size, and family density—upon children's behavior and upon their subsequent behavior and adjustment as adults. Yet one obvious and seemingly important variable—parental age—has received surprisingly little attention as far as research is concerned, although, certainly, it has been the subject of lay and professional opinion for years.

There are references in the literature to the relation of parental age to children's intelligence [e.g., Thurstone and Jenkins (11)] and to a host of physical and nonbehavioral characteristics such as congenital malformations, mongolism, longevity, and sex ratio of offspring [e.g., Bleyer (2) and Strong (10)]. Schooler (8), in speculating on the relation between birth order and mental illness, has noted that apparent birth-order effects may be an artifact of the relation between advancing mother's age and the offspring's disorder; and Warren (12) has commented that mother's age may be a significant factor in the origin of birth-order effects. Some British investigators [see Goodman (4)] have investigated the relation between maternal age and incidence of mental disorder in adults. Goodman found that hospitalized schizophrenics showed a slightly higher maternal age distribution than was to be expected on the basis of census data on all births, and Norton (4) found that a group of outpatient neurotic adults tended to have older mothers than did a group of controls.

E. V. Cowdry, in an introduction to a monograph devoted to the subject of parental age and characteristics of the offspring, states what is probably a commonly held notion: ". . . the impact on the developing personalities of the offspring exercised by elderly, instead of by young, parents may perhaps force different kinds of adjustments . . ." [see Strong (10, p. 453)]. Yet neither this research monograph nor others pursue the relation between parental age

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 21, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

and the kinds of adjustment children make or the types of problems they manifest.

A serendipitous finding by the present authors in a previous study (9) suggested that parental age is related to the type of problem manifested by children seen in a psychiatric clinic. Both the mothers and fathers of children with "personality problems" (i.e., problems associated with the over-inhibition of impulses), as defined by a problem checklist [see Peterson (7)], were significantly older than mothers and fathers of children with "conduct problems" (problems associated with the expression of impulses). The authors felt that the apparent limitations of the previous study—small sample size, as well as heterogeneity of the children's age within the sample—underscored the need for further investigation.

The present study sought to substantiate this finding in a large sample of elementary-school children, in which the age of the child could be held constant. Unfortunately, only mothers' ages were available, so the effect of fathers' ages could not be investigated further.

B. METHOD

Available from another study (6) were factor scores on Peterson's problem checklist for 518 second-grade children (252 boys and 266 girls) in the Cedar Rapids, Iowa, public-school system; also available on 341 of these children was information on the mother's age at the time of the child's birth and on the child's ordinal position in the family.[1]

As in the authors' previous study (9), only children of decided problem types were used as subjects. Since there are sex differences in the type of problems manifested by children (7), the data were analyzed separately for boys and girls. Thus, 33 boys whose scores on the conduct factor were above the median[2] of the large sample of second-grade boys and whose scores on the personality factor were below the sample median were selected for the boys' conduct-problem group. Similarly, 28 boys whose personality-factor scores were above the sample median and whose conduct-factor scores were below the median were selected for the boys' personality-problem group. In like manner, 21 girls were selected for the girls' conduct-problem group, and 45 for the girls' personality-problem group. Mean ages of the groups of children were as follows: boys' conduct-problem group, 8.2 years; boys' personality-problem group, 7.9 years; girls' conduct-problem group, 8.2 years; and girls' personality-problem group, 8.0 years.

---

[1] The problem checklists were completed by the children's teachers in the spring of the school year; hence the teachers had had ample opportunity to become acquainted with the children's problems. Factor scores were computed by unweighted summation over the 15 variables or problems making up each of the two factors on the problem checklist. The mothers' ages and the number of older and younger siblings were reported on a questionnaire mailed to the mothers.

[2] The median conduct-factor score for the larger sample of 252 second-grade boys was 2.8, and the median personality-factor score was 3.5. The median conduct-factor score for the sample of 266 second-grade girls was .9, and the median personality-factor score was 2.7. Medians were used since the distributions of factor scores were J-shaped. Such a distribution is not unexpected, since the majority of children do not present any significant problems in the school setting.

C. RESULTS

Mean ages of the mothers of the children in the different problem-type groups were compared by means of *t* tests, as shown in Table 1.

TABLE 1
COMPARISONS OF MEAN AGES OF MOTHERS (AT CHILD'S BIRTH) BY CHILD'S PROBLEM TYPE

Problem groups	*N*	Mothers' ages Mean	Mothers' ages *SD*	*t*
Boys				
Conduct problem	33	24.8	5.1	1.869*
Personality problem	28	27.3	5.4	
Girls				
Conduct problem	21	25.0	5.4	0.704
Personality problem	45	25.9	4.2	
Combined (boys and girls)				
Conduct problem	54	24.8	5.2	1.752*
Personality problem	73	26.4	4.8	

* $p < .05$ (p values are for a one-tailed test of significance).

To determine if problem type is related to birth order, the proportion of first-born and later-born children in the conduct-problem and personality-problem groups (boys and girls combined) were compared by means of chi square test. The resulting chi square value of .508 ($df = 1$) is not significant.

D. DISCUSSION

These results substantiate the previous finding that mothers of children with personality problems are significantly older than mothers of children with conduct problems. In this study it is apparent that the difference is not a function of the child's age, since the personality-problem groups were not older than the conduct-problem groups. It does appear that the sex of the child is an important variable, since the mother's-age variable is significant in the comparison between the boys' groups but not between the girls' groups. It also appears that type of problem manifested by second-grade children is not related to their birth order. It should be noted that the difference in age of the mothers of the two problem groups is only a difference of a few years—not a difference of a generation.

Since one would assume that children with personality problems (as defined by Peterson) are more likely to become neurotic adults than are children with conduct problems, the present results seem consistent with Norton's finding [see Goodman (4, p. 203)] that adult neurotic patients tended to have older mothers than did controls.

The specific basis for the relationship between mother's age and the child's problem type is still a matter of speculation. One's first assumption might be that older mothers have attitudes about child-rearing different from those of younger mothers. Indeed, many studies have demonstrated a relation between mother's age and professed attitudes about child-rearing. Curiously enough, however, meaningful relations between measures of parental attitudes and children's behavior problems have been demonstrated only rarely [see Becker and Krug (1)]. Perhaps parental age is a mediating variable that is obscured by the measurement error of attitude scales.

It is also possible that the few years' difference between the early and late 20s is important in terms of biological factors in maternity. In any event, while it is obvious that behavior is the result of a complex interaction of the organism and its environment, it would appear that the simple matter of parental age may partially determine the type of adjustment children make.

This finding may have implications in regard to the shape of psychopathology to come: as the age at which women bear children decreases, as it has in recent years (3), one may expect to see an increase in the number of conduct problems in psychiatric clinics and guidance centers. It may also have some implication in regard to the adjustment of adopted children, since adoptive parents are typically seven to eight years older than their biological counterparts when their first child comes (5).

E. SUMMARY

The present study was undertaken to substantiate a previous finding that parental age is associated with the type of adjustment that children make. From a sample of 518 second-grade public-school children, 127 who were rated by their teachers as having decided personality and conduct problems were selected. It was found that mothers of children with personality problems were significantly older than mothers of conduct-problem children.

REFERENCES

1. BECKER, W. C., & KRUG, R. S. The Parental Attitude Research Instrument—A research review. *Child Devel.,* 1965, **36**, 329-365.
2. BLEYER, A. Childbearing Before and After Thirty-Five. New York: Vantage Press, 1958.
3. GLICK, P. G. American Families. New York: Wiley, 1957.
4. GOODMAN, N. Relation between maternal age at parturition and incidence of mental disorder in the offspring. *Brit. J. Prev. & Soc. Med.,* 1957, **11**, 203-213.
5. KIRK, H. D. Shared Fate. New York: Free Press of Glencoe, 1964.
6. NORRIS, A. S., & ROWLEY, V. N. Capillary morphology, prenatal factors, and children's behavior problems. Unpublished study, University of Iowa, Iowa City, Iowa, 1963.
7. PETERSON, D. R. Behavior problems of middle childhood. *J. Consult. Psychol.,* 1961, **25**, 205-209.
8. SCHOOLER, C. Birth order and schizophrenia. *Arch. Genet. Psychiat.,* 1961, **4**, 91-97.
9. STONE, F. B., & ROWLEY, V. N. Children's behavior problems and parental attitudes. *J. Genet. Psychol.,* 1965, **107**, 281-287.
10. STRONG, L. C., *Ed.* Parent age and characteristics of the offspring. *Ann. N. Y. Acad. Sci.,* 1954, **57**, 451-613.
11. THURSTONE, L. L., & JENKINS, R. L. Order of Birth, Parent Age, and Intelligence. Chicago: Univ. Chicago Press, 1931.
12. WARREN, J. R. Birth order and social behavior. *Psychol. Bull.,* 1966, **65**, 38-49.

Child Psychiatry Service
State Psychopathic Hospital
500 Newton Road
The University of Iowa
Iowa City, Iowa 52241

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 235-247.

THE RELATIONSHIP OF CONFIDENCE TO COMPETENCE AND IMITATION BEHAVIOR*[1]

American Institutes for Research, Washington, D. C.

HERBERT HOROWITZ

A. INTRODUCTION

An ever growing body of research in interpersonal influence suggests a strong relationship between, on the one hand, conformity and imitativeness (usually studied as dependent variables) and, on the other hand, the competence characteristics displayed by the subject and by relevant others in the situation (usually studied as independent variables). Studies that show such an effect can be grouped into one of three categories: (*a*) those in which the *subject's competence* at the task was varied: e.g., Kelman (14) and Mausner (17); (*b*) studies in which the *competence of others* was varied: e.g., Mausner (16, 18), Rosenbaum and Tucker (22), Kanareff and Lanzetta (11, 13), and Kidd and Campbell (15); and (*c*) studies that involve simultaneous manipulations of *subject's competence and others' competence:* e.g., Patel and Gordon (20), Croner and Willis (5), Samelson (24), DiVesta (6), Kanareff and Lanzetta (12), Rosenbaum, Chalmers, and Horne (21), and Mausner and Block (19). Taken collectively, these studies warrant the following conclusion: factors that lead the subject to believe that he is incompetent or that others are more competent than he are associated with heightened susceptibility to social influence.

One way of accounting for such an effect is to hypothesize that the competence-affecting procedures used in these experiments lead to variations in the *confidence* the subject has in his own ability *vis-à-vis* the confidence he has in the ability of others. In turn, these variations in confidence level serve to mediate the observed levels of conformity and imitativeness. In addition to

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 21, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This study is based on a Doctoral dissertation submitted to the Graduate School of the University of Wisconsin. The study was conducted while the author held a U. S. Public Health Service Social Psychology Traineeship. The author wishes to thank the members of his research committee, J. C. Gilchrist (Chairman), B. Martin, and V. L. Allen for their guidance, and V. J. Polidora for his assistance in designing the apparatus.

A brief version of this report was presented at the annual meetings of the American Psychological Association in Chicago, Ilinois, 1965.

the studies cited above, the literature contains other studies that suggest even more directly such a relationship between confidence and conforming behavior. For example, Wiener (25) reported that "uncertainty of judgment, a response attribute," and not the stimulus attribute of ambiguity was responsible for conformity behavior. In agreement with this finding, Fisher, Williams, and Lubin (10) reported a fairly high rank-order correlation between conformity and degree of self-certainty regarding the stimulus. Similarly, Rosenberg (23) found that the amount of conformity that individuals displayed was (*a*) positively related to the degree of error they attributed to their own judgments, and (*b*) negatively related to the degree of error attributed to others.

It must be recognized initially that in using the concept of confidence to account for the relationship between imitativeness and competence, confidence is viewed as a *cognitive state*. Thus, any expression of confidence obtained is seen as the end product of a series of cognitive processing steps. These processes involve two kinds of cognitions: (*a*) the perceptions and thoughts the individual has about his own abilities and the abilities of relevant others for making correct decisions, and (*b*) the subjective probabilities that the individual develops from such thoughts and perceptions concerning the likelihood that others will be correct more often than he himself in the future.

To predict the occurrence of imitative behavior, it is insufficient to know only the amount of competence displayed by the subject or by others: i.e., the models, whom he imitates. It is also necessary to know how much confidence the subject has in the model *vis-à-vis* his self-confidence. In other words, the incorporation of the construct, "confidence," in an explanation of imitation requires a multidimensional conceptualization in which the subject's initial *and* subsequent confidence in his model as well as in himself can be specified.

With this conceptual framework the present study sought to relate certain cognitive states—namely, the confidence an individual has in himself and in the advice of others—to (*a*) the stimuli from which those cognitive states are developed: i.e., the competence of the advice and the individual's own competence, and (*b*) the behavioral consequences of those cognitions: i.e., the degree of imitativeness shown.

B. METHOD

1. *Subjects*

Eighty female college students, mostly from introductory psychology courses, served as subjects, for which they received experimental participation points. Since the subjects were to be used in pairs, efforts were made to insure that each pair consisted of strangers. In a handful of cases, however, the subjects did know each other slightly.

2. *Task and Apparatus*

The task was to make a series of predictions of an unknown person's reactions to a set of attitude-opinion statements. The person whose reactions were to be predicted by the subjects had purportedly given his reaction—i.e., "agree" or "disagree"—to each of these statements sometime prior to the experiment. The subject's task was to predict whether the choice for a given statement had been "agree" or "disagree." The statements were taken from the Bales Value Profile (1), Borgatta's Multi-Level Personality Inventory (3), the Edwards Personal Preference Schedule (8), or were composed by the experimenter. Each statement was presented on a separate four- by six-inch index card.

On the table in front of each subject was the response panel, which contained seven jeweled lamps and two momentary push-button switches. Six of these lights were arranged in three rows of two lights each. The top row was labeled "Your Partner's Choice," the second "Correct Answer," and the third "Your Choice." The left-hand column of three lights was designated "Agree," and the right column "Disagree." Below the third row and centrally located was a seventh lamp, marked "Your Turn," which served as the subject's signal to respond.

The apparatus was designed to perform three functions. The first was to program an initial 20-trial *induction* phase according to any one of four reinforcement schedules: low or high rate of reinforcement for the subject and, independently, low or high rate of reinforcement for the partner (the model during the experimental period). The second function of the apparatus was to program the series of model's choices and knowledge of results during the second, *experimental,* period consisting of 160 trials. The third operation of the apparatus was to prepare an Esterline-Angus event record of subjects' responses, programmed models' choices, and programmed feedback. This record was used for scoring imitativeness of the subject.

3. *Procedure*

Each subject worked in a separate room. During the initial period, consisting of 20 trials, each subject made her prediction before her "partner's" prediction for that trial was revealed. Thus, the subject's responses were entirely independent of her partner's responses during this *competence induction period.* After each prediction, the partner's response was presented, as was the alleged "correct" answer (the "choice actually made" by the person being predicted). Thus, the subject was able to determine whether she and her partner had been correct on that prediction. The partner's responses were programmed in such a way that the partner (model-to-be) appeared either high or low in competence: i.e., her responses were either 70 per cent or 30 per cent correct. Simultaneously, and *independently,* the subject's own competence was also made to be either high or low, that is, her responses were either 70 per cent or 30 per cent correct. The four possible combinations of model-competence and subject-competence established the four basic conditions of the experiment to each of which were assigned 20 subjects (Total $N = 80$).

After the 20 predictions that constituted the induction phase (having gotten either six or 14 correct and her apparatus-simulated partner having gotten either six or 14 correct), each subject received additional (pretaped) instructions summarizing her own and her (simulated) partner's performance. Since each pair of subjects was assigned to the same treatment, each subject in the pair heard the same evaluation of her own and her partner's performance. These instructions also explained that the purpose of the remaining part of the investigation was to determine if people's predictive abilities *improve* as a result of prior knowledge of another person's predictions.

After these instructions, all subjects went through the 160 trials composing the *experimental period.* The sequence of events during each experimental trial was: model's choice followed by the subject's reading of the stimulus statement followed by the subject's response and then feedback. Thus, the subject was unable to make a judgment before the model's response was known to her. The model's choices purportedly originated with the partner who was working in an adjacent room. These partner responses were, in fact, programmed electrically by the apparatus and were identical for all subjects. Regardless of how competent the partner-model had been during the initial 20-trial induction period, during the first half of the experimental period the model was made to appear incompetent by making her choices correct only 30 per cent of the time. During the last 80 trials the model's competence was *reversed* by making her correct 70 per cent of the time. These two levels of model competence appeared in the same sequence for all subjects and comprised two levels of the third factor of the experiment, *subsequent model competence.*

In summary, the experimental design was a factorial 2 × 2 × 2: (*a*) Initially High Competent Subject *vs.* Initially Low Competent Subject (HCS *vs.* LCS); (*b*) Initially High Competent Model *vs.* Initially Low Competent Model (HCM *vs.* LCM); and (*c*) Low Subsequent Model Competence *vs.*

High Subsequent Model Competence (designated First Half *vs.* Second Half). Each subject appeared in only *one* combination of the first two factors, and at *both* levels of the last factor, in the same order.

4. *Dependent Variables*

Two dependent variables were examined. The first was imitativeness, measured by counting the number of times the subject's response matched the response given by the model in each block of 10 trials. In order to determine the relationship between confidence and imitativeness, a series of measurements was made of the subject's self-confidence and her confidence-in-the-model. The following questions were used in operationalizing these confidence determinations:

> Using one of the numbers from zero to 10 indicate:
> 1. How many of the next 10 predictions you will get correct?
> 2. How many of the next 10 predictions your partner will get correct?

The first presentation of these questions was after the 10th experimental-period trial, and thereafter these questions were completed again after each additional 20 trials, a total of eight times.

At the end of the experiment, subjects provided answers to a series of post-experimental questions designed to ascertain, among other things, their confidence in themselves and in the model.

C. RESULTS

The results can be subsumed under three headings: (*a*) the effects of the independent variables on the subject's confidence in herself and in the model; (*b*) the effects of these variables on imitativeness; and (*c*) the relationship between imitation behavior and confidence.

1. *Confidence Ratings*

Three kinds of analyses were performed in order to assess the effects of the treatments on the subject's confidence. The first method was to analyze the periodic ratings of the subject's self-confidence and, separately, the periodic ratings of her confidence in the model.

Table 1 summarizes the analysis of variance of the first set of self-confidence ratings, those obtained after the 10th experimental trial. It can be seen that *both* factors, subject's initial competence and model's initial competence, significantly affected the confidence that the subject expressed in her own ability. Subjects who had been made initially incompetent at the task expressed significantly *less* confidence in their own ability than those who had been made initially competent. In addition, and surprisingly, those subjects paired with a partner who was initially *competent* expressed *more* self-confidence than those paired with a partner *low in competence*. That is, the model's competence affected the *subject's self-confidence*. Both these effects can be traced to those subjects who had been made high in initial ability and who were paired with a partner who was also high in ability. These subjects expressed the greatest self-confidence of all four combinations of model competence (HCM-HCS, HCM-LCS, LCM-HCS, and LCM-LCS).

TABLE 1

ANALYSIS OF VARIANCE OF RATINGS OF SELF-CONFIDENCE OBTAINED DURING *Initial* MEASUREMENT (AFTER TENTH EXPERIMENTAL TRIAL)

Source	*df*	*MS*	*F*
Model's initial competence (A)	1	6.05	4.88*
Subject's initial competence (B)	1	11.25	9.07**
A × B	1	4.05	3.27
Subjects within (A × B)	76	1.24	
Total	79		

* $p < .05$.
** $p < .01$.

A separate analysis was performed on the first set of ratings of confidence-in-the-model (after the 10th experimental-period trial). This analysis showed that confidence-in-the-model was affected by the partner-model's initial competence, though not by the subject's initial competence. Subjects paired with an initially competent model had more confidence in the model than those with an initially incompetent model ($F = 29.66$, $p < .001$ for 1,76 *df*).

Confidence-in-the-model was affected not only by *induction period* variations in model's competence, but also by the variation in subsequent model's competence. This is shown by an analysis of variance of the eight measures of confidence-in-the-model that were taken throughout the experiment (see Table 2). In this analysis the first four measures of confidence (obtained during the first 80 trials) were pooled and compared with the last four (second 80 trials). The results of Table 2 are in agreement with those of the previous analysis in showing again that the model's initial competence affected to a reliable degree the confidence in the model's ability. However, confidence in the model was also influenced by the change in competence occurring between the two halves of the experiment. Regardless of the model's initial competence, the mean rating of confidence in the model was significantly higher during the second half when she was competent than during the first half when she was incompetent. Although the interaction of initial model competence with subsequent model competence did not yield a significant *F* ratio, a Duncan Multiple Range Test (7) revealed that the increase in confidence-in-the-model between halves was greater for subjects paired with an initially incompetent partner than for those with a competent partner. In other words, they showed a relatively large increase in confidence-in-the-model when her competence was switched.

TABLE 2
ANALYSIS OF VARIANCE OF RATINGS OF CONFIDENCE-IN-THE-MODEL IN EACH 80-TRIAL EXPERIMENTAL PERIOD

Source	*df*	*MS*	*F*
Model's initial competence (A)	1	316.40	14.28***
Subject's initial competence (B)	1	5.25	< 1.00
A × B	1	15.01	< 1.00
Subjects within (A × B)	76	22.16	3.16***
Subsequent model competence (C)	1	120.75	17.23***
A × C	1	23.27	3.32
B × C	1	.06	< 1.00
A × B × C	1	1.80	< 1.00
Subjects within (A × B) × C	76	7.01	
Total	159		

Note: Analysis in this table is based on four scores pooled within each 80-trial period.
*** $p < .001$.

A second kind of analysis of confidence ratings was based on the difference between the subject's confidence in herself (c_s) and her confidence in the model (c_m) at each of the eight times of measurement. The analysis of variance of these $c_s - c_m$ scores[2] (pooling across the four measures in each half of the experiment) is shown in Table 3. This analysis reveals that when the subject is initially competent at the task, predictably, there is a greater difference between her self-confidence and her confidence in her partner, the model, than when the subject is initially incompetent. Furthermore, these difference scores are greater when the model is low in ability (first half) than when she shows much ability (second half). The last finding of interest in this analysis is the nearly significant interaction between subject's and model's initial competence level. A Duncan Multiple Range Test (alpha = .01) on the four means constituting this interaction revealed that the HCM-LCS group gave significantly lower $c_s - c_m$ scores than any of the other three groups, who did not differ from each other. In other words, confidence-in-self relative to the other person was affected not only by the subject's own competence, but by her own competence *vis-à-vis* the model's competence.

---

[2] These $c_s - c_m$ scores were algebraically summated. Since some $c_s - c_m$ scores had negative values, the largest negative value was used as a constant that was added to all scores, thereby making them positive. The analysis of variance in Table 3 and the correlations reported are based on these positive scores.

TABLE 3
ANALYSIS OF VARIANCE OF $c_s - c_m$ SCORES IN EACH 80-TRIAL EXPERIMENTAL PERIOD

Source	*df*	*MS*	*F*
Model's initial competence (A)	1	39.80	2.07
Subject's initial competence (B)	1	93.05	4.85*
A × B	1	73.10	3.81[a]
Subjects within (A × B)	76	19.20	2.27***
Subsequent model competence (C)	1	714.05	84.60***
A × C	1	14.59	1.73
B × C	1	21.00	2.49
A × B × C	1	3.41	< 1.00
Subjects within (A × B) × C	76	8.44	
Total	159		

Note: Analysis in this table is based on four scores pooled within each 80-trial period.
[a] $p = .06$.
* $p < .05$.
*** $p < .001$.

The last method of determining the effects of the experimental manipulations on the subject's confidence in herself and in the model is based on responses to two questions from the postexperimental questionnaire. In the first of these, subjects were asked to estimate the percentage of correct responses made *by the model* during the 160 experimental trials. An analysis of variance was performed on these estimates, and this analysis showed that a significantly higher mean estimate was given by subjects in the HCM condition than by subjects in the LCM treatment (56.75 per cent *vs.* 45.54 per cent, $F = 14.10$, $p < .01$). In a second question, subjects were asked whether they felt that they had more, less, or the same amount of ability as the partner (the model) for making such predictions. Disregarding those 45 subjects who answered "same," of the remaining 18 subjects in the LCM conditions, half responded "more" and half responded "less." In contrast, nearly all (14 of the 17) HCM subjects responded "less," while the other three indicated "more" ($\chi^2 = 4.06$, $p < .05$ for 1 *df*). In other words, even at the end of the experiment, subjects who had been exposed to an initially competent model were more likely to feel less able than the model when compared to subjects paired with an initially incompetent model.

From these results it seems clear that the manipulations of the subject's and the model's competence affected (*a*) the subject's perceptions of the model's ability during the experiment and (*b*) her appraisal of her own ability relative to the model's ability.

2. *Imitation Behavior*

A three-way analysis of variance of imitation scores was computed and is summarized in Table 4. These results reveal that there is *no* main effect due either to the model's or the subject's initial competence. However, there is a very clearcut difference as a function of variations in *subsequent* model competence. Imitativeness *was* significantly greater when the model was competent than when she was incompetent, a result that matched the results obtained by Kanareff and Lanzetta (12, 13) and by Rosenbaum, Chalmers, and Horne (21).

TABLE 4
ANALYSIS OF VARIANCE OF IMITATION RESPONSES

Source	*df*	*MS*	*F*
Model's initial competence (A)	1	127.41	1.44
Subject's initial competence (B)	1	142.09	1.60
A × B	1	23.42	< 1.00
Subjects within (A × B)	76	88.63	3.73***
Subsequent model competence (C)	1	956.58	40.21***
A × C	1	91.17	3.83[a]
B × C	1	2.31	< 1.00
A × B × C	1	48.47	2.04
Subjects within (A × B) × C	76	23.79	
Total	159		

[a] $p = .06$.
*** $p < .001$.

In addition to the effect of changes in the model's experimental-period competence, there is also evidence that this difference is, in turn, affected by the level of competence displayed *initially* by the model. This evidence stems from the marginally significant *F* ratio for the interaction of initial and subsequent model competence. The four means comprising this interaction were subjected to a Duncan Multiple Range Test (alpha = .01). The results of this test indicated that a *greater* difference in imitativeness occurs between the two levels of subsequent model competence (i.e., the two halves) when the model was initially incompetent as compared to those conditions in which the model was initially competent. This pattern of interaction reflects a similar one described above for *confidence* in the model.

3. *Relationship Between Confidence and Imitativeness*

The main thesis of the present investigation is that differences in competence, between oneself and another, mediate changes in self-confidence and that these confidence levels, in turn, are related to the amount of interpersonal dependence, imitativeness, exhibited. To test this proposition, a series of product-moment correlations were computed between $c_s - c_m$ scores and the number of imitations occurring in the subsequent block of 10 trials. Since each individual was confidence-questioned eight times, each subject provided eight pairs of scores. These eight pairs of scores were correlated separately within each of the four conditions.

The following correlations were obtained in each condition: HCM-HCS = —.130; HCM-LCS = —.167; LCM-HCS = —.168; and LCM-LCS = —.148. Though the magnitude of these correlations is, in absolute terms, quite small, all four correlations are significantly different from zero (at $p \leqslant .05$). Furthermore, all correlations have negative signs, indicating that the greater the difference between the subject's confidence in herself and her confidence in the model, the less likely she is to imitate the model.

D. Discussion

In his 1954 "theory of social comparison processes," Festinger (9) presented a series of propositions and supporting data concerning the processes and circumstances under which individuals compare their own opinions and abilities with those of others. The present experiment extends Festinger's ideas by placing the subject in a situation in which social comparisons were unavoidable. Thus, in the present experimental paradigm, the individual has both nonsocial (knowledge of results) and social (forced self- and model-evaluations) bases for determining how successful her adjustment is to the experimental environment.

The decision-making environment in which the subjects performed was structured in such a way that success was directly contingent on their perceiving correctly and then utilizing (either by opposing or imitating) the advice from another: i.e., the model's choices. The data presented, while matching the findings of other studies (e.g., Kanareff and Lanzetta (12, 13) and Rosenbaum, Chalmers, and Horne (21)) in showing again that competence manipulations affect the level of imitativeness, go beyond the earlier studies in showing that competence manipulations lead not only to differences in imitativeness but also to alterations in self-confidence and in confidence-in-others. Furthermore, the correlations reported, between imitativeness and differences in self-confidence and confidence-in-the-model, show that a reliable relationship exists between the overt behavior that is usually examined in studies of conformity and imitation, and another dimension, confidence, that is cognitive in nature and serves a mediating function.

The fact that the confidence scores used in these correlations were obtained *prior* to the imitation scores with which they were correlated provides a major source of support for the contention that confidence does, in fact, mediate imitativeness. Nevertheless, it is necessary to list two important qualifications to the conclusion that imitativeness is a reflection of confidence-in-self *vis-à-vis* confidence-in-other. First, the subjects in this experiment were *forced* to make these social comparisons. It may be that there is no correlation between confidence and imitativeness when social comparisons are avoidable.

The second qualification is based on the magnitude of the correlations that were found between $c_s - c_m$ scores and imitation scores. The conceptual scheme presented above views the subject as a hyper-rational decision-making unit, one that can calculate subjective probabilities accurately and then base decisions on the most adaptive strategy, either opposition or imitation. However, the size of the obtained correlations shows that the subjects did not act as rationally as the theory predicts. The fact that the correlations account for such a small portion of the total variance of imitation scores (approximately 2 per cent) indicates that other factors are operating. Clearly, the bulk of variance in imitation scores can be accounted for by the combined effect of situational and organismic factors that were not systematically examined in this study but which a large body of research has shown to be important (2, 4). Undoubtedly, additional research that varied situational factors (like the nature of the decision task, the nature of the nonsocial feedback, the nature and size of incentive) and organismic factors (like the sex of the subject, personality characteristics, age, actual ability at the task) would show effects on both the magnitude of $c_s - c_m$ scores and also on the strength of association between such scores and imitation behavior. Overriding these obvious limitations on generalizability, the present data nevertheless warrant the conclusion that a temporal-sequential relationship does indeed exist between confidence parameters and at least one kind of informational social influence process.

E. SUMMARY

The present study tested the hypothesis that *the confidence a subject has in himself vis-à-vis another person* mediates the often-replicated negative association between the person's relative ability and his imitativeness. Eighty female college students predicted an unknown person's responses to a long series of attitude-opinion statements. Prior to the experimental period, each subject was made to appear either competent or incompetent at the task and, simultaneously and independently, her partner was also made to appear either high or low in competence at the task. After this competence induction period, each subject made a long series of similar predictions, always after the partner's

(the model's) choice had been revealed to her. The model's competence was varied again during the experimental period.

The results showed that the manipulations of the subject's initial ability and the model's initial and subsequent abilities significantly affected (*a*) the subject's confidence in herself and her confidence in the model, and (*b*) her imitativeness during the experiment. In addition to these effects, support for the major hypothesis was presented in the form of significant negative correlations between the level of imitation behavior and the difference between the subject's self-confidence and her confidence-in-the-model.

REFERENCES

1. BALES, R. F. The Value Profile, Form E. Unpublished, undated manuscript.
2. BLAKE, R. B., & MOUTON, J. S. Conformity, resistance and conversion. In I. A. Berg & B. M. Bass (Eds.), *Conformity and Deviation*. New York: Harper, 1961.
3. BORGATTA, E. F. The Multi-level Personality Inventory. Unpublished manuscript, 1961.
4. CAMPBELL, D. T. Conformity in psychology's theories of acquired behavioral dispositions. In I. A. Berg & B. M. Bass (Eds.), *Conformity and Deviation*. New York: Harper, 1961.
5. CRONER, M., & WILLIS, R. Perceived differences in task competence and asymmetry of dyadic influence. *J. Abn. & Soc. Psychol.,* 1961, **62**, 705-708.
6. DIVESTA, F. J. Effects of confidence and motivation on susceptibility to informational social influence. *J. Abn. & Soc. Psychol.,* 1959, **59**, 204-209.
7. DUNCAN, D. B. Multiple range and multiple F tests. *Biometrics,* 1955, **11**, 1-16.
8. EDWARDS, A. L. Edwards Personal Preference Schedule. New York: Psychol. Corp., 1953.
9. FESTINGER, L. A theory of social comparison processes. *Hum. Rel.,* 1954, **7**, 117-140.
10. FISHER, S., WILLIAMS, H. L., & LUBIN, A. Personal predictors of susceptibility to social influence. *Amer. Psychol.,* 1957, **12**, 360 (Abstract).
11. KANAREFF, V. T., & LANZETTA, J. T. The acquisition of imitation and opposition responses under two conditions of instruction-induced set. *J. Exper. Psychol.,* 1958, **56**, 516-528.
12. ————. Effects of success-failure experience and probability of reinforcement upon the acquisition and extinction of an imitative response. *Psychol. Rep.,* 1960, **7**, 151-166.
13. ————. The effects of task definition and probability of reinforcement upon the acquisition and extinction of imitative responses. *J. Exper. Psychol.,* 1960, **60**, 340-348.
14. KELMAN, H. C. Effects of success and failure on "suggestibility" in the autokinetic situation. *J. Abn. & Soc. Psychol.,* 1950, **45**, 267-285.
15. KIDD, J. S., & CAMPBELL, D. T. Conformity to groups as a function of group success. *J. Abn. & Soc. Psychol.,* 1955, **51**, 390-393.
16. MAUSNER, B. Studies in social interaction: III. Effect of variation in one partner's prestige on the interaction of observer pairs. *J. Appl. Psychol.,* 1953, **37**, 391-393.
17. ————. The effects of prior reinforcement on the interaction of observer pairs. *J. Abn. & Soc. Psychol.,* 1954, **49**, 65-68.
18. ————. The effects of one partner's success in a relevant task on the interaction of observer pairs. *J. Abn. & Soc. Psychol.,* 1954, **49**, 557-560.

19. MAUSNER, B., & BLOCK, B. L. A study of the additivity of variables affecting social interaction. *J. Abn. & Soc. Psychol.*, 1957, **54**, 250-256.

20. PATEL, A. S., & GORDON, J. E. Some personal and situational determinants of yielding to influence. *J. Abn. & Soc. Psychol.*, 1960, **61**, 411-418.

21. ROSENBAUM, M. E., CHALMERS, D. K., & HORNE, W. C. Effects of success and failure and the competence of the model on the acquisition and reversal of matching behavior. *J. of Psychol.*, 1962, **54**, 251-258.

22. ROSENBAUM, M. E., & TUCKER, I. F. The competence of the model and the learning of imitation and nonimitation. *J. Exper. Psychol.*, 1962, **63**, 183-190.

23. ROSENBERG, L. A. Conformity as a function of confidence in self and confidence in partner. *Hum. Rel.*, 1963, **16**, 131-139.

24. SAMELSON, F. Conformity behavior under two conditions of conflict in the cognitive field. *J. Abn. & Soc. Psychol.*, 1957, **55**, 181-187.

25. WIENER, M. Uncertainty of judgment as a determinant of conformity behavior. *Amer. Psychol.*, 1956, **11**, 407 (Abstract).

American Institutes for Research
8555 Sixteenth Street
Silver Spring, Maryland 20910

Published as a separate and in *The Journal of Psychology*, 1966, 63, 249-260.

CAN INTERVIEW INTERACTION MEASURES BE TAKEN FROM TAPE RECORDINGS?[*][1]

University of Oregon Medical School

Arthur N. Wiens, Donald A. Molde, David C. Holman,
and Joseph D. Matarazzo

A. Introduction

Research on the noncontent (and content) aspects of interview behavior has accelerated rapidly within the past two decades. Numerous investigators, in addition to our own group, have shown interest in the use of objective measurement devices to collect and accumulate data on such noncontent variables as speech and silence characteristics in interview situations. One of the first to produce such a device was Chapple (2). His Interaction Chronograph was designed to record electromechanically and thus quantify the noncontent (temporal) interaction processes occurring in an interview. Another investigator (3) used a pen recorder from which she also derived Chapple-type information regarding the length of speech utterances and silences. Kasl and Mahl (5) slightly modified this system to meet their own special needs. Our own group used the Chapple Interaction Chronograph in similar research for over a decade (6), and more recently we developed a fully electronic paper-tape punch model which we have called the Interaction Recorder (12).

All of the above investigators have employed an observer who has recorded the interview "live" to derive their speech and silence measures. Some investigators have taken the observer out of the system. Verzeano and Finesinger (11) used a microphone which allowed them to record the interview data automatically. Hargreaves and Starkweather (4) modified this approach to form their Duration Tabulator which automatically records and analyzes the distributions of speech unit durations. Recently Cassotta, Feldstein, &

---

[*] Received in the Editorial Office, Provincetown, Massachusetts, on April 25, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported by Public Health Service Research Grant No. NU 00133 from the Division of Nursing, Bureau of State Services—Community Health, and from Research Training Grant No. 5T-T5-GM59 from the National Institute of General Medical Sciences, National Institutes of Health. Thanks are also due to Jean Boyle, Director of the School of Nursing, and to her staff for their help in the recruitment of some of the subjects.

Jaffe (1) have developed an Automatic Vocal Transaction Analyzer (AVTA) which translates speech recorded into two channel tape recordings directly onto IBM cards. Our experience has led us to conclude that automatic devices that record directly without the mediation of a human observer distort the final temporal data collected. Although expensive recording devices are available, only a human observer can adequately assess (usually by verbal context), for example, whether a pause in the middle of an ongoing statement is just that or the end of one communicative speech unit.

The appearance of such a variety of recording devices indicates the growing interest among investigators in recording speech and silence durations in a variety of types of interview settings (10). However, the relative immobility of such bulky machinery (e.g., our Interaction Recorder or Jaffe's AVTA) precludes its transportation to those everyday interview situations that occur beyond reach of the microphone cords or the building in which the bulky recording device is housed. This difficulty, coupled with the impossibility of always having an observer available, prompted our own group to see whether the speech and silence measurements we have used for a decade needed to be made "live" at the time of the actual interview or whether, hopefully, as is now often done in clinical or other types of research practice, the interview could be tape-recorded and the speech and silence analyses made later from the tape recording. Thus, the main purpose of the present study was to determine whether interaction data retrieved from a tape recording of an interview were comparable to data obtained from a "live" observation of the same interview. The ability to tape-record an interview in a far removed location for later analysis by the now extant relatively fixed recorder-computer systems would greatly increase the range and flexibility of interview research programs.

The design of the present study involved the following steps. First, an experienced observer recorded "live" 20 partially standardized interviews. A tape recording was simultaneously made of each interview. Second, two novice observers, independent of each other and independent of knowledge of the experienced observer's results with each interview, themselves recorded each of the same 20 interviews from the tape recordings alone. Third, the "live" observer, using the tapes alone, rescored the last 10 interviews several weeks after the original interviews had taken place. This third step made it possible to compare directly and effectively the results obtained from the tapes alone with the results obtained by the "live" observer. The research design also permitted the two novice "tape observers" to be compared with each other, and also with the "live" observer.

B. METHOD

1. *Apparatus*

The length of each interviewer's and interviewee's speech and silence units was recorded on the Interaction Recorder (12). This is an electronic device which time-records on paper tape an account of the time when either person in an interview is speaking or is silent. These recordings are in the form of binary code which is acceptable to a Burroughs E101 computer. An interview is recorded "live" through a one-way mirror by an observer who depresses either the interviewer or interviewee key, depending upon who is talking. Both keys are depressed if both participants are talking at the same time. Each key is released at the completion of an utterance, providing for a sequential analysis of the interview. A Bell & Howell tape recorder was used to tape-record simultaneously the "live" interview.

2. *Observers*

Three different observers were employed in this study. The first (ANW) was an experienced research psychologist who had recorded hundreds of interviews of this type. He recorded each interview "live" from behind a one-way mirror. Sound was transmitted through earphones which were connected to a microphone embedded in the ceiling of the interview room. Thus the first observer had available both *visual* and *auditory* cues from the interview in progress. The second and third observers were a second- and a fourth-year medical student, both of whom had seen very few standardized interviews and had, themselves, scored none. A day or two after each interview they both recorded that interview from a tape recording which had been made at the time of the "live" interview. One of these observers (DAM) was also the interviewer in 19 of the interviews. We included this feature of having the interviewer later record from his own tapes because we felt that in many instances in which research of this type is being carried on it will be the interviewer, himself, who later will do the analyses from his own tapes. We did, however, include in our design a second observer (DCH) who had no prior knowledge of the interviews on the tapes, in order to test the influence of such prior knowledge on later "observer" recordings. Thus the second observer (DCH) had neither been present at the interview nor had he seen it. Therefore, he was scoring it completely "cold." As part of the design of this study the "live" observer (ANW) was instructed that nonverbal forms of communication (i.e., head nodding) were not to be recorded, since it would be impossible for the two other observers to derive these forms of communication from the tape recording. We believe this to present little problem, since in recording interview interactions we have found, with our interviewers at least, head nodding and other forms of nonverbal communication to constitute a negligible fraction of the total.

3. *Subjects*

The interviewees in this study were selected without design from the range of persons available on a medical campus, such as ours, and they came from widely varied backgrounds. Ten of them were female, first-year student nurses. Two were medical secretaries; two were medical students; three were general medical patients; one was a psychiatric patient; one was a urology patient; and one was a postpartum obstetrics patient.

4. *Interviewers*

The interviewer for this study was a fourth-year medical student (DAM). He was relatively inexperienced at conducting a partially standardized research interview (see below) having previously interviewed about five subjects. The interviewer for the first interview only was an experienced faculty psychologist, who conducted the first interview as a sample for the medical student interviewer, who then conducted the next 19 interviews. Later analysis revealed that this one interview did not affect the results in any way and thus it was included with the subsequent 19 interviews to complete the total of 20.

Except for some minor rules which we included in order to allow us to relate the findings of the present study to our earlier and ongoing studies, the interview procedure utilized with all 20 *S*s was a typical nondirective interview. The partial standardization utilized the following rules: all interviews were approximately 30 minutes in length; the interviewer talked approximately five seconds per utterance; the interviewer responded to the interviewee's last statement within one second; the interviewer used no nonverbal methods of communication (i.e., head nodding, etc.); and the interviewer used open-ended questions. The content of these questions (i.e., nursing, present job, family, hobbies, schooling, present illness, etc.) produced a relaxed interview atmosphere.

5. *Variables Studied*

In keeping with our earlier research [summarized in Matarazzo, Wiens, & Saslow (6)] the following speech and silence variables were recorded and analyzed for this study:

a. Interviewee mean speech duration. This variable includes the total amount of time (in seconds) the interviewee spoke divided by his total number of speech units. A speech unit, as defined by us, is an utterance separated at either end by two silence periods—one silence following the other participant's last comment (i.e., the speaker's latency), and the second silence following the speaker's own comment and preceding the listener's next comment (i.e., the listener's latency). Pauses for breathing, for choosing words, for reflection, etc., are included in the speech unit when the *context* clearly suggests that the speaker has not yet completed that utterance. (An example of one utterance or unit is: "Yes, I can tell you about . . . [pause] . . . my father. He was a . . . [pause] . . . how can I say it . . . [pause] . . . a man of many talents. You would . . . [pause] . . . have . . . [pause] . . . liked him.") However, pauses (again determined by context) which precede the introduction of new ideas or thoughts by the same individual, *without an intervening comment by the other interview participant,* signal the onset of a new speech unit. (For example: "It's true that I've . . . [pause] . . . enjoyed hunting for most of my life." . . . [pause three to four seconds] . . . "Speaking of hunting, I'm reminded of a favorite dog I had several years ago.") The context clearly suggested that one unit ended with "life" and a second, or new unit, began with "Speaking." The reader is also referred to Rogers (9, pp. 265-437) and Wolberg (13, pp. 688-780) for numerous examples of typical single interview units as scored by us. We have found observer reliability for these units and the other variables here described to be unusually high (8).

b. Interviewee mean speech latency. This variable includes the total latency time divided by the number of units of interviewee latency. A latency is that period of silence separating two *different* speech units. Latencies occur either between exchanges of the interview participants (i.e., at the time one terminates and the other begins an utterance) or between two consecutive but clearly separate and different speech units spoken by one individual.

c. Interviewer mean speech duration and interviewer mean speech latency. This variable is the same as those for the interviewee, with change of the word interviewer for interviewee above.

C. Results and Discussion

The important findings of this study are shown in Tables 1 and 2. For the 20 interviews, ANW, scoring the interviews "live," obtained an average mean duration of interviewee utterance of 51.5 seconds; while DAM, who conducted the interviews, and later scored them from the tapes, obtained a mean of 44.4

TABLE 1

SPEECH AND SILENCE DURATION MEASUREMENTS BY DIFFERENT OBSERVERS ON THE SAME 20 INTERVIEWS

	Observers			Correlations between observers		
Variables	ANW	DAM	DCH	ANW-DAM	ANW-DCH	DAM-DCH
Interviewee duration of utterance (seconds)						
Mean	51.5	44.4	48.2	.98***	.99***	.97***
Sigma	39.6	31.5	38.9			
Range	11.7 to 154.8	11.9 to 116.4	12.4 to 155.0			
Interviewee duration of latency (seconds)						
Mean	.77	.85	.86	.97***	.99***	.98***
Sigma	.74	.72	.74			
Range	.26 to 3.75	.26 to 3.56	.29 to 3.75			
Interviewer duration of utterance (seconds)						
Mean	5.1	5.1	5.0	.95***	.95***	.92***
Sigma	.6	.6	.5			
Range	4.2 to 6.2	4.3 to 6.1	4.3 to 6.2			
Interviewer duration of latency (seconds)						
Mean	.71	.55	.75	.58**	—.12	.47*
Sigma	.23	.16	.21			
Range	.36 to 1.25	.24 to .92	.33 to 1.10			

* Significant at .05 level of probability.
** Significant at .01 level of probability.
*** Significant at .001 level of probability.

TABLE 2
Speech and Silence Duration Measurements by Different Observers on the Same 10 Interviews

	Means among observers				
Variables	ANW-"live"	ANW-"tape"	DAM	DCH	Correlation between observers
Interviewee duration of utterance (seconds)					
					ANW Live-ANW Tape = 1.00***
Mean	34.4	35.1	30.3	30.2	ANW Live-DAM Tape = .97***
					ANW Live-DCH Tape = .99***
Sigma	12.0	12.5	10.8	10.0	ANW Tape-DAM Tape = .96***
					ANW Tape-DCH Tape = .99***
Range	19.7 to 57.7	19.9 to 59.0	15.9 to 49.2	18.7 to 48.9	DAM Tape-DCH Tape = .98***
Interviewee duration of latency (seconds)					
					ANW Live-ANW Tape = .97***
Mean	.66	.73	.83	.77	ANW Live-DAM Tape = .94***
					ANW Live-DCH Tape = .97***
Sigma	.28	.36	.39	.34	ANW Tape-DAM Tape = .97***
					ANW Tape-DCH Tape = .96***
Range	.26 to 1.16	.30 to 1.40	.26 to 1.54	.29 to 1.33	DAM Tape-DCH Tape = .98***
Interviewer duration of utterance (seconds)					
					ANW Live-ANW Tape = .92***
Mean	5.0	5.2	4.9	5.0	ANW Live-DAM Tape = .96***
					ANW Live-DCH Tape = .94***
Sigma	.5	.5	.5	.6	ANW Tape-DAM Tape = .97***
					ANW Tape-DCH Tape = .97***
Range	4.15 to 5.94	4.61 to 6.16	4.33 to 5.86	4.34 to 6.16	DAM Tape-DCH Tape = .98***
Interviewer duration of latency (seconds)					
					ANW Live-ANW Tape = .30
Mean	.59	.67	.53	.87	ANW Live-DAM Tape = .45
					ANW Live-DCH Tape = .20
Sigma	.14	.17	.11	.18	ANW Tape-DAM Tape = .71*
					ANW Tape-DCH Tape = .22
Range	.36 to .79	.47 to .93	.37 to .75	.57 to 1.10	DAM Tape-DCH Tape = .62

* Significant at .05 level of probability.
*** Significant at .001 level of probability.

seconds; and DCH, scoring them "blind," obtained a mean of 48.2 seconds for the same 20 interviewees. Analysis by Pearson *r* revealed that ANW correlated .98 with DAM and .99 with DCH; and DAM correlated .97 with DCH. It is clear from the very high correlations obtained for this variable (and the next two variables) that each observer is essentially interchangeable with the others on any of these three variables. That is, each of the "tape" observers correlated very highly with the "live" observer as well as with each other.

However, in Table 1 there is a spread of 7.1 seconds between the interviewees' mean duration of utterance as recorded by the three observers. Likewise, there are spreads of .09 second between interviewees' mean latency, and .1 second between the interviewer's mean duration of utterance. Statistically significant *t* tests across these correlated means were obtained for interviewee duration of utterance (ANW-DAM, $t = 2.96$, p of .01; and ANW-DCH, $t = 3.51$, p of .01), and for interviewee latency (ANW-DCH, $t = 3.00$, p of .01). Although these three of the total of nine *t* tests which were computed reached statistical significance, one might ask what is the practical significance of these slight but statistically significant differences. Inspection of our raw data (not presented in either table) indicates that the significant *t* tests obtained are a reflection of individual differences in the scoring techniques of the three observers. DAM and DCH tended to obtain values which differed from ANW's values in a constant way (i.e., shorter or longer measurements depending on the variable). For instance, in 15 of 20 instances, both DAM and DCH obtained smaller values for interviewee mean duration of utterance than did ANW. With interviewee latency, ANW scored this variable shorter in 12 of 20 instances than did DCH. That this kind of idiosyncratic behavior should become evident when individuals make a series of precise relatively subjective judgments—such as deciding the exact moment when an utterance has been completed, especially when the intervening latency between one utterance and the next is often very small, a fraction of one second—is not surprising. When one is dealing with hundredths of a second or even several seconds one is clearly dealing with the reaction times of individual observers, a classic problem in Psychology.

To date, none of the previous research of our group can attribute any practical significance to the difference of seven seconds between interviewee duration of utterance shown in Table 1. For instance, in Matarazzo, Wiens, & Saslow (6), inspection of the figures on pp. 199 and 205, which depict interviewee duration of utterance in various studies already completed, will make clear that a seven-second variation in any of those means would have negligible effects in altering the overall results. The same is true for the .1 second difference found in the observer's recording of interviewer duration of utterance. As shown in Table 3 of Matarazzo, Wiens, Saslow, Allen, & Weitman (7, p. 112) a variability of .09 second in interviewee latency values also would have no practical effect on the final results obtained to date.

The differences seen between the means of the first three variables shown in Table 1 might well be similar to those that could be expected to occur when the same observer repeats the same measurements of the same interview. Our research design allowed us to test this possibility, since ANW scored 10 interviews "live" and later scored them again from their tape recordings. Table 2 presents data from the last 10 interviews of the series. This table compares not only the values obtained by the "live" observer (ANW) with those of the two "tape" observers but also includes the values obtained when the "live" observer scored the tapes of the last 10 interviews several weeks after the original interviews took place, thereby allowing some measure of the reliability of a single observer repeating similar measurements over time and under "live" and "tape" conditions. Inspection of the means and ranges obtained by ANW-"live" and ANW-"tape," as shown in Table 2, discloses minimal variation between them. In measuring interviewee mean duration of utterance, ANW-"live" recorded 34.4 seconds, ANW-"tape" 35.1 seconds. The range for this variable is quite similar: 19.7 to 57.7 seconds for ANW-"live," 19.9 to 59.0 seconds for ANW-"tape." Interviewee latency shows for ANW-"live": mean = .66, range = .26 to 1.16 seconds; for ANW-"tape": mean = .73, range = .30 to 1.40 seconds. Similarly, interviewer duration of utterance also shows minimal variation. Comparable results are shown for the two other tape observers. The small actual difference in means coupled with the very high correlations (none less than .92) for the first three variables shown in Table 2 makes it doubtful that any practical significance can be attributed to these discrepancies for the same reasons listed above for the discrepancies noted in Table 1.

The last variable, interviewer latency, presents a more difficult problem in this study. Table 1 shows that two of the three latency correlations reached significance (ANW-DAM, $r = .58$, p of .01; DAM-DCH, $r = .47$, p of .05), but that they are considerably smaller than the correlations obtained with the other three variables. Significant t tests were obtained between ANW-DAM ($t = 3.90$, p of .01) and DAM-DCH ($t = 5.00$, p of .001). Though two of the means are very similar (ANW, .71; DCH, .75), the lack of correlation between these two observers ($r = -.12$) indicates that this similarity is more fortuitous than real. Likewise, Table 2 shows only moderate agreement among observers concerning this same mean interviewer latency variable for the last 10 interviews.

It is possible, however, that for this variable closer agreement is easier to achieve between two "tape" observers than between a "live" observer and a "tape" observer. Table 2 reveals that the three correlations between live and tape recordings were .30, .45, and .20, whereas for tape to tape recordings the correlations were .71 (*p* of .05), .22, and .62 (*p* of .06). A relevant factor may be that in the absence of visual cues, the task of deciding from the tapes alone just when the interviewee has finished an utterance and the interviewer is about to enter the conversation is a difficult one. Hargreaves and Starkweather (4) report a similar problem. In comparing two different observations of the same recording of a lecturer, these authors were dissatisfied with the differences between the two and attributed these differences to the inadequacy of strictly auditory cues in allowing anticipation of stops and starts in the lecturer's speech.

A further complicating factor regarding the interviewer's latency measurement is our rule which directed the interviewer to respond to the interviewee's last comment in less than a second. This produced a small numerical value for the interviewer latency and a narrow range of scores (.36 to 1.25 seconds). At this level, reaction time is such a prominent factor that any good comparability in recording this variable will be difficult to achieve, both between "live" and "tape" observers, and tape observers only, without the use of some additional recording aids.

Two possible recording aids are the following: the first involves the use of a constant lag. The observer tries to be just a "moment" behind the interaction for *all* utterances and silences. (A "moment," in time dimensions, is roughly .5 second.) Hopefully, the constant lag will always allow the observer an instant longer to decide whether the speaker has finished his utterance. A danger should be avoided, however, in that it is easy to develop a rhythm whereby "extra" silence is introduced into short latency periods. A second method of overcoming some of the scoring difficulties with this last variable is to listen to the tape recording prior to recording it with the Interaction Recorder. Notes can be taken on exchanges which will be difficult to score, thus allowing some anticipation and forewarning.

When our overall findings are reviewed with the four variables under study, it is apparent that comparable data were obtained among three different observer-recorders (whether they recorded "live" or from a "tape,") of an interview which had been conducted earlier. This comparability was evident in the high correlations and similar mean values for the different observers when they recorded the interviewees' duration of utterance and latency and the interviewer's duration of utterance. Even when the more difficult variable, interviewer latency was dealt with—with numerical values of less than one second and undoubtedly involving individual differences in observer reaction time—gross comparability for all practical purposes was evident in mean values that ranged only from .55 to .75 second and in correlations between observers that were significant in two out of three cases (Table 1). Nevertheless, reliability is less for this variable and it may be necessary to utilize some recording aids that will enhance the comparability of data recorded for this variable and thereby decrease to a considerable degree the variance among observers, "live" or "tape."

We conclude from our findings in this study that it is feasible to collect data initially by tape-recording interviews and subsequently re-recording them on the Interaction Recorder or similar devices for analysis of speech and silence interaction characteristics. Thus, by not being tied to our relatively fixed recorder-computer system the range and flexibility of our own interview research program and that of other investigators can be increased. For example, we can now design a study to compare the interview characteristics of different occupational groups or subspecialty groups within a given occupational or professional group and go into that work setting to obtain data. Or, research collaboration between our own group studying noncontent interview characteristics and another group, perhaps physically removed, studying its content could be established with confidence that data collected by tape recording would be comparable to data collected live (and necessitating physical presence and proximity).

REFERENCES

1. CASSOTTA, L., FELDSTEIN, S., & JAFFE, J. AVTA: A device for automatic vocal transaction analysis. *J. Exper. Anal. Behav.*, 1964, **7**, 99-104.
2. CHAPPLE, E. D. The Interaction Chronograph: Its evolution and present application. *Personnel,* 1949, **25**, 295-307.
3. GOLDMAN-EISLER, F. Speech analysis and mental process. *Lang. & Speech,* 1958, **1**, 59-75.
4. HARGREAVES, W. A., & STARKWEATHER, J. A. Collection of temporal data with the Duration Tabulator. *J. Exper. Anal. Behav.*, 1959, **2**, 179-183.
5. KASL, S. V., & MAHL, G. F. A simple device for obtaining certain verbal activity measures during interviews. *J. Abn. & Soc. Psychol.*, 1956, **53**, 388-390.
6. MATARAZZO, J. D., WIENS, A. N., & SASLOW, G. Studies of interview speech behavior. In L. Krasner & L. P. Ullmann (Eds.), *Research in Behavior Modification: New Developments and Clinical Implications.* New York: Holt, Rinehart & Winston, 1965. Pp. 179-210.
7. MATARAZZO, J. D., WIENS, A. N., SASLOW, G., ALLEN, B. V., & WEITMAN, M. Interviewer mm-hmm and interviewee speech durations. *Psychother.: Theory, Res. & Practice,* 1964, **1**, 109-114.

8. PHILLIPS, J. S., MATARAZZO, J. D., MATARAZZO, R. G., & SASLOW, G. Observer reliability of interaction patterns during interviews. *J. Consult. Psychol.*, 1957, **21**, 269-275.

9. ROGERS, C. R. Counseling and Psychotherapy. Boston: Houghton Mifflin, 1942.

10. SIEGMAN, A. W., & POPE, B. Effects of question specificity and anxiety-producing messages on verbal fluency in the initial interview. *J. Personal. & Soc. Psychol.*, 1965, **2**, 522-530.

11. VERZEANO, M., & FINESINGER, J. E. An automatic analyzer for the study of speech in interaction and in free association. *Science*, 1949, **2845**, 45-46.

12. WIENS, A. N., MATARAZZO, J. D., & SASLOW, G. The Interaction Recorder: An electronic punched paper tape unit for recording speech behavior during interviews. *J. Clin. Psychol.*, 1965, **21**, 142-145.

13. WOLBERG, L. R. The Technique of Psychotherapy. New York: Grune & Stratton, 1954.

Department of Medical Psychology
University of Oregon Medical School
3181 S. W. Sam Jackson Park Road
Portland, Oregon 97201

Published as a separate and in *The Journal of Psychology*, 1966, 63, 261-268.

ON NONVERBAL INTERACTION*

Psychology Department, San Francisco State College

JOSEPH LUFT

A. INTRODUCTION

This is a study of nonverbal interaction, an attempt at finding ways of measuring and understanding what takes place affectively between two persons in a nonverbal task-free situation.

Dyadic interaction is clearly one of the fundamental processes in clinical and social psychology, and personality theorists reveal their basic differences when they attempt to explain what takes place in a dyad and why. Some account must be made of large classes of variables, such as what the subject brings to the situation, the context in which interaction occurs (6), what the persons perceive (7), the conscious and unconscious needs involved, and the variables in the behavioral interchange itself. The problem of the researcher is to hold down the number of variables so that some sense can be made of the behavior observed. However, this is not quite as simple as it sounds for, as Cronbach (3) remarked after examining the results of many studies in this field, the findings are "interesting, significant and exasperatingly inconsistent."

B. CONFRONTING STRANGERS

One clue about the significance of meeting persons for the first time comes from Charles H. Cooley (2), who wrote in his *Human Nature and the Social Order,*

> There is a vague excitement of the social self more general than any particular emotion or sentiment. Thus the mere presence of people, a "sense of other persons" and an awareness of their observation, often causes a vague discomfort, doubt and tension. One feels that there is a social image of himself lurking about, and not knowing what it is he is obscurely alarmed. Many people, perhaps most, feel more or less agitation and embarrassment under the observation of strangers, and for some even sitting in the same room with unfamiliar or uncongenial people is harassing and exhausting. It is well known, for instance, that a visit from a stranger would often cost Darwin his night's sleep, and many similar examples could be collected from the records of men of letters. At this

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 28, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

> point, however, it is evident that we approach the borders of mental pathology (2, p. 207).

C. Problems and Method

Basically what we did in the present study was to bring together two subjects who did not know each other and to instruct them to sit across from each other in a room so that they were about seven to nine feet apart. The first 14 subjects were freshman and sophomore women who served as volunteers. They were told not to talk or to communicate in any way. After five minutes, each subject noted on a simple 20-point like-dislike scale how she felt toward the other person. This was repeated for a total of three ratings covering 15 minutes. Each subject was then removed to another room for a short interview. The purpose of the interview was to learn more about what was taking place in the interaction from a subjective point of view. A few days later, the experiment was repeated until a total of five meetings had taken place. During the meeting, two experimenters acted as observers, noting through a one-way screen such things as gestures, motions, glances, avoided looks, and other facial or bodily behavior.

We found from our observations and our interviews following each 15-minute meeting that just about every subject experienced considerable anxiety and discomfort. This is reminiscent of studies on sensory deprivation, in which *S*s are placed in settings so that the usual background of auditory, visual, and tactile stimulation is drastically restricted. Our situation, by analogy, might thus be called "social stimulus deprivation."

1. *The Rating Dilemma*

In our experiment, the subject actually faces a difficult and unfamiliar problem. Although she has undoubtedly been in nonverbal face-to-face situations with a stranger, the subject could at least keep her feelings to herself. But now she is forced to make up her mind and to reveal her personal reactions. Furthermore, she can see that the other *S* was rating her, although she cannot see the form or the ratings. Someone is obviously going to see her ratings and she, the subject, will be making herself vulnerable. In addition, *S* has an obligation to herself, to her own integrity, to express how she really feels.

And yet, on what can she base her reaction? It is inevitable that her rating will reflect, at least in part, her own predilections, her own experience with other people. But, in addition, she has before her the other unique individual with that individual's ongoing, though muted, behavior. At this point, it is postulated that true interpersonal testing takes place and only part of this testing may be done consciously. For example, how does the other subject respond to her and to the small nonverbal cues that she sends out? Is there an attempt at understanding her inquiring glance, or is it coldly ignored? Does the other subject display postural cues of tension, indicating some distress at confronting her? Does the other subject grow increasingly comfortable, indicating some kind of acceptance; or does the other subject treat her as if she were a thing and did not exist? These and many other kinds of readily discernible behavioral exchanges appear to take place. If *S* is open enough and free enough to take in and to respond to these minor communications and is able to take a quasi-public stand about how she feels, then we may get an authentic record of her reactions to this situation. If she is not free or not sensitive to the present social situation, then we may get either gross projections and distortions or we may get some kind of denial and evasion.

A few *S*s showed such denial and evasion by giving rigidly high favorable ratings. We found that such persons were high scorers on the Berkeley F Scale.

2. *Changes in Feeling over Time*

One of our main questions, then, was what happens to affect ratings over time? The results of interaction in the first three dyads are shown in graphic form (see Figure 1).

3. *Conflict Between Data Perceived and Data Given by Experimenter*

The data suggest that *S*s tend to converge and to rise in their ratings. In order to determine whether *S*s could be influenced to lower their like ratings, we arranged three new dyads as before. Then, after *S*s had an opportunity to experience two sessions, we told each *S* that the other member of the dyad was rating her lower than she was rating the other member, regardless of the actual ratings. We gave this information during the usual preliminaries before each session. In two of the three dyads, there was an immediate drop in the ratings followed by a slow rise to approximately previous positions. In the third dyad, no change occurred, but this was the only pair where fairly strong dislike ratings had been given independently by each before the feedback session.

In two dyads, we gave positive feedback before their third meeting and their scores either rose or remained at a fairly high level.

A few of the *S*s suspected that the information given them was spurious, but most apparently accepted it at face value.

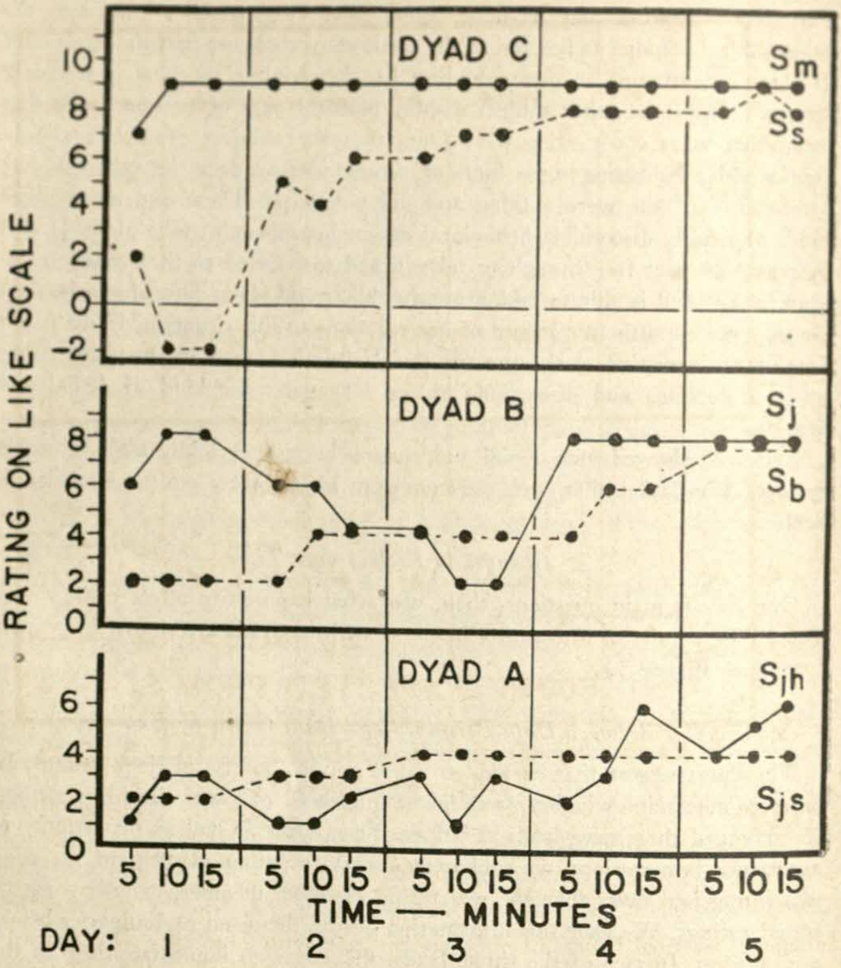


FIGURE 1
THREE PATTERNS OF DYADIC INTERACTION

4. *Ratings Given Compared to Ratings Expected*

Another question we considered was this: can *S*s judge the ratings given them by the other person in the dyad? Almost without exception, we found that *S*s expected ratings that were very close to the ratings they themselves gave the other member (see Figure 2). It would seem from this that people

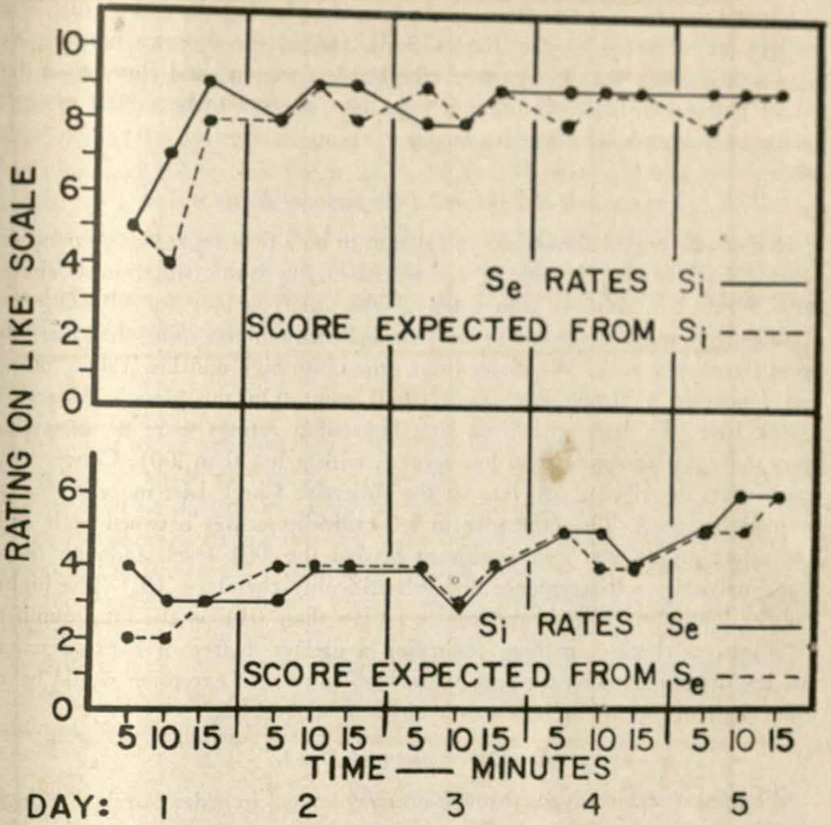


FIGURE 2
RATINGS GIVEN COMPARED TO RATINGS EXPECTED

anticipate reactions from others commensurate with how they feel toward others; and, frequently, these expectations are way out of line with what the other *S* actually feels, at least in the ambiguity of the experimental situation. Tagiuri (9) finds the same tendencies in his "Social Preference and Its Perception" (9, p. 323).

5. *Anxiety and the Perceived Physical Distance Between Subjects*

Because of the ambiguity and stress of the experimental situation, we speculated that when *S*s are removed from the room in which they confront each other, they will recollect the distance between *S*s in a distorted way.

We found that in six out of seven dyads, the person having the greater manifest anxiety, as measured by the Taylor Scale, judged the distance between *S*s significantly closer than her partner who had less anxiety, and closer than the actual physical distance. In other words, there appears to be a kind of generalized claustrophobic effect for anxious persons in stressful interaction with others.

6. *Self-Regard and Interpersonal Ratings*

How do *S*s regard themselves in relation to how they regard or respond to others? Each of the 36 subjects was placed in the dyadic situation as above with a research assistant who acted as the other partner for all subjects. Ratings were made at the end of the first and fifth minutes using the same 20-point like-dislike scale. We divided our group into high and low raters, using five (moderately strong like) as a cutoff point. On the Gough Adjective Check List (5), high scorers on first impression ratings were persons who were more self-accepting than low scorers, with p less than .001. Conversely, low scorers described themselves on the Adjective Check List more critically than high scorers. The difference in self-criticality scores between high and low affect groups was also significant beyond the .001 level. (Gough finds that empirically self-acceptance and self-criticality correlate .15.) The highs and the lows also differed significantly (p less than .01) on the total number of adjectives checked, perhaps indicating a greater degree of expressiveness and less defensiveness on the part of the high scorer. An exception would have to be made of the rigidly high scorer of the affect ratings.

7. *Observation of Dyads*

We tried observing dyads through one-way screens in order to relate ratings to the phenomena observed. The meaning of eye contacts appears to be especially important. This is one point at which folklore and literature and clinical observations find some agreement. Simmel (8) has stated, "The union and interaction of individuals is based upon mutual glances. This is perhaps the most direct and purest reciprocity which exists anywhere" (8, p. 358). Although it was possible to observe gestures and glances, etc., there is a need for increased accuracy and reliability; and perhaps motion picture photography is most appropriate here.

D. A CONCLUDING NOTE

These experiments are, as mentioned, frankly exploratory. What is significant is that the procedure lends itself to the study of central and molar phenomena in interpersonal behavior. A number of preliminary findings are supported by other studies, such as those in the ONR symposium on *Person Perception and Interpersonal Behavior* reported by Tagiuri and Petrullo (10). Goffman's basic ideas in his book, *The Presentation of Self in Everyday Life* (4), are highly relevant to some of the questions we have raised in this study. Also, personality theorists appear to be shifting away from preoccupation with the person *per se,* as Blake and Mouton (1) observe in their review of research in personality. "Thus it becomes necessary to develop a way of concept formation which deals with persons and situations simultaneously" (1, p. 203). In the relatively simple interaction situation here proposed, we may have a way of studying interpersonal behavior in a setting that permits such simultaneous considerations.

E. Summary

The impact of strangers on each other in a nonverbal arrangement was measured in terms of changes in like-dislike ratings. Hypotheses concerning the effect of personality variables on the interaction were examined. Fourteen college students in seven dyads met in a series of five 15-minute confrontations. Also, 36 *S*s interacted nonverbally with a standard stimulus subject. The results suggest that attitude toward self is positively related to attitudes toward strangers and that affect ratings expected from others follow closely the rating given to others. High-anxiety *S*s tended to recall the physical distance between the dyad partners as closer than did the low-anxiety partner, and as closer than the actual distance.

References

1. Blake, R., & Mouton, J. S. Personality. *Ann. Rev. Psychol.,* 1959, **10**, 203-232.
2. Cooley, C. Human Nature and the Social Order. New York: Scribners, 1902.
3. Cronbach, L. J. Proposals leading to the analytic treatment of social perception scores. In R. Tagiuri & L. Petrullo (Eds.), *Person Perception and Interpersonal Behavior*. Stanford, Calif.: Stanford Univ. Press, 1958. Pp. 353-379.
4. Goffman, E. The Presentation of Self in Everyday Life. Garden City, N. Y.: Doubleday, 1959.
5. Gough, H. Adjective Check List. Palo Alto, Calif.: Consult. Psychol. Press, 1961.
6. Luft, J. Interaction and projection. *J. Project. Techn.,* 1953, **17**, 489-492.
7. ————. Monetary value and the perception of persons. *J. Soc. Psychol.,* 1957, **46**, 245-251.
8. Simmel, G. Sociology of the senses: Visual interaction. In R. E. Park (Ed.), *Introduction to the Science of Sociology*. Chicago: Univ. Chicago Press, 1921. Pp. 356-360.
9. Tagiuri, R. Social preference and its perception. In R. Tagiuri & L. Petrullo (Eds.), *Person Perception and Interpersonal Behavior*. Stanford, Calif.: Stanford Univ. Press, 1958. Pp. 316-336.

10. TAGIURI, R., & PETRULLO, L., *Eds.* Person Perception and Interpersonal Behavior. Stanford, Calif.: Stanford Univ. Press, 1958.

Psychology Department
San Francisco State College
1600 Holloway Avenue
San Francisco, California 94132

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 269-274.

DIFFERENTIAL EFFECTS OF COFFEE ON SPEED AND POWER TESTS*

Psychological Department, University of Düsseldorf, Germany

G. A. LIENERT AND H. P. HUBER[1]

A. PROBLEM

Considerable information has been gathered concerning the effects of coffee on psychological test performance. Positive as well as negative results have been reported [see Eichler (4), Gilliland and Nelson (6), and De Boor (2, p. 126-130)]. Careful study of these reports suggests that positive effects occur most often in performance on tests requiring speed. Until now the speed-power characteristics of the tests applied have not been controlled when determining the influence of coffee on the test performance. The following study was carried out to examine the question of whether the effect of coffee on performance of speed tests is different from its effect on performance of power tests.

B. METHOD

1. *Subjects*

Two randomly selected groups of *S*s, an experimental group and a control group, were used. Each group consisted of 50 male and female students from the Liberal Arts and Sciences Departments of the University of Marburg. The students ranged in age from 20 to 32 years.

2. *Tests*

Each *S* was given two standardized tests in an individual session: (*a*) the Intelligenzstrukturtest (IST) by Amthauer (1), a battery of nine more or less speeded power tests, and (*b*) the Konzentrationsleistungstest (KLT) by Düker and Lienert (3), a typical speed test consisting of 250 relatively easy arithmetical items of the form $8 - 3 + 7 = ?$

---

* Accepted by Joseph Zubin of the Editorial Board, received in the Editorial Office, Provincetown, Massachusetts, on April 29, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] The experiment was performed by L. Kangieser, junior student at the Psychological Department of the University of Marburg (Germany), under supervision of the senior author.

3. *Treatment*

Fifteen minutes prior to testing, *Ss* of the experimental group were given a cup of strong coffee (Nescafé) containing .3 grams of caffeine. *Ss* of the control group did not receive any drink.

C. Hypotheses and Predictions

It was hypothesized that speed of test performance would be increased by caffeine, whereas power of performance would remain unchanged or be decreased. The theoretical implication was that the dynamic factors of intelligence should be positively influenced; the cognitive factors on the other hand should be influenced either not at all or perhaps negatively. From this hypothesis the following predictions were derived:

1. Performance in the typical speed test (KLT) will increase under coffee.

2. Performance in the more or less speeded power tests of the IST battery will increase or decrease depending upon their degree of speeding.

The degree of speeding of a single test of the IST battery was to be measured by its correlation with the typical speed test KLT.

D. Results

With respect to the first prediction, the standard score means and variances ($\mu = \sigma^2 = 100$) of the speed test (KLT) presented in Table 1 were obtained. These data indicate that the coffee group performed better than the control group. The difference between means is significant at the 1 per cent level of confidence. Thus Prediction 1 is confirmed.

Results concerning the means and variances of the standard scores of the partly speeded IST subtests ($\mu = \sigma^2 = 100$) are given in Table 2. According to Table 2, performance in three of the nine subtests increased (numerical reasoning, RA, inductive reasoning, ZR, and spatial reasoning, FA); whereas performance in two other subtests (general information, SE, and verbal comprehension, GE) decreased.

To decide whether the subtests positively influenced by coffee require more speed than the subtests not positively influenced, product-moment-correlations

TABLE 1
Means and Variances for Both Groups on the Speed Test

KLT	Experimental group	Control group	*N*
μ	116.1	102.3	50
σ^2	122.9	106.9	50

TABLE 2
MEAN STANDARD SCORES AND VARIANCES FOR BOTH GROUPS ON THE NINE IST SUBTESTS

IST subtest	Experimental group ($N = 50$)		Control group ($N = 50$)	
	μ	σ^2	μ	σ^2
SE = General information	108.0	41.3	116.6	38.6
WA = Verbal classification	110.2	65.1	108.9	55.1
AN = Semantic analogies	111.0	54.3	112.3	47.1
GE = Verbal comprehensions	110.0	40.4	115.0	80.8
ME = Associate memory	107.9	58.3	108.8	51.9
RA = Numerical reasoning	108.9	107.1	107.4	102.4
ZR = Inductive reasoning	112.5	85.3	110.2	150.1
FA = Spatial relations	110.5	55.1	106.8	134.6
WÜ = Spatial orientation	107.8	117.4	106.7	114.3

between the KLT and each of the nine IST subtests were calculated and tested against the null hypothesis according to Table 3. Two of the three tests favored by coffee were found to be positively correlated with the typical speed tests, while the correlations of those tests not positively influenced did not differ significantly from zero. If all of the nine subtests are considered, a high positive rank correlation of rho = .82 was found between the speed criterion for the subtests and gains in performance under coffee. Thus, Prediction 2 must be also accepted as true.

E. DISCUSSION

The fact that only speeded tests were favored by coffee may be explained in terms of Thurstone's multiple factor theory of intelligence (10). The speed factor or the speed factors—at least one of which has been found in factorial studies [see Eysenck (5, p. 152) and Lord (9)]—should be of increased importance under coffee. The positive or negative effects of coffee upon particular tests may be the result of gains and losses, respective of this speed factor.

As yet unmentioned is the partial result that the variance of performance on some of the IST subtests (GE, ZR, and FA) was significantly smaller for the coffee group than for the control group (see Table 2). Obviously, coffee tends to reduce individual differences in specified intellectual abilities, as measured by the individual IST subtests. In terms of the multiple factor theory of intelligence, the effect may be explained tentatively by supposing different factorial test patterns in high and low scoring *S*s, a phenomenon that has been shown to exist [see Lienert (7)]. If, for example, the inductive reasoning test (ZR), given to a group of low-scoring *S*s, gains more in speed factor loading under coffee than the same test given to a group of high-

TABLE 3

SUBTEST INTERCORRELATIONS (PEARSONIAN *r*'S) AND MEAN GAIN (IN STANDARD SCORE UNITS) UNDER COFFEE

IST subtest	SE	WA	AN	GE	ME	RA	ZR	FA	WÜ	N
Average correlation (*r*) with the KLT from both groups	.17	.22	—.04	—.16	.23	.45	.39	.29	.24	2 × 50
Mean gain under coffee	—3.6	1.3	—0.7	—5.0	—0.9	1.5	2.4	3.7	1.1	50

scoring *S*s, test performance of the low-scoring *S*s will increase more than test performance of the high-scoring *S*s, thus reducing variance of test performance. It is not established that the variance-reducing effect of coffee has anything to do with a hypothesized differential speed factor pattern. The problem should be examined empirically by separately factorizing high- and low-scoring groups of *S*s with and without administration of coffee.

Also open to question is the validity of our speed criterion. The correlations between the KLT and the IST subtests may not be due exclusively to common variance in the speed factor, but also to the common variance in a numerical factor. The latter may have unduly increased the correlation between the arithmetical KLT and the numerical tests ZR and RA. However, this objection may be rejected by considering a further speed criterion defined by Lienert and Ebel (8) and calculated for each of the nine IST subtests. With regard to this criterion, the two numerical tests (RA and ZR) were shown to be highly speeded just as with the original criterion. Rank correlation between the two criteria was found to be .72.

Finally, it should be stressed that the present investigation was not planned or performed to test the effects of caffeine nor—strictly speaking—to test the effects of coffee, but to test the effects of drinking coffee on performance in speed and power tests. Therefore the criticism that a placebo medication was not given to the *S*s of the control group is not relevant. It is clear that the intake of coffee produced some uncontrolled sets, and therefore it cannot be stated that the changes in test performance, observed in this experiment, are due solely to the effect of coffee. They will be at least partly due to the unlike motivational level of the two groups compared. So our conclusions are not far from being practically significant.

F. SUMMARY

To test the hypothesis of differential effects of coffee on speed and power tests, two groups of 50 *S*s each were given 10 standardized tests, one of which was a typical speed test and the other nine of which were partly speeded power tests. Before testing, the *S*s of the experimental group were given a cup of Nescafé containing .3 grams of caffeine, while the *S*s of the control group did not receive any drink.

As predicted from prior empirical evidence and from theoretical considerations, mean performances of the speed tests and the mostly speeded power tests were increased significantly for the coffee group, whereas mean per-

formance in the less or nonspeeded power tests was either not affected or reduced.

Results were interpreted in terms of a multiple-factor theory of intelligence by changes in speed factor loading of the tests under caffeine.

REFERENCES

1. AMTHAUER, R. Der Intelligenzstrukturtest (IST) (2nd ed.). Göttingen: Hogrefe, 1955.
2. DE BOOR, W. Pharmakopsychologie und Psychopathologie. Berlin: Springer, 1956.
3. DÜKER, H., & LIENERT, G. A. Konzentrationsleistungstest (KLT). Göttingen: Hogrefe, 1959.
4. EICHLER, O. Kaffee und Coffein. Berlin: Springer, 1938.
5. EYSENCK, H. J., *et al.* Dimensions of Personality. London: Routledge, 1947.
6. GILLILAND, A. R., & NELSON, D. The effects of coffee on certain mental and physiological functions. *J. Gen. Psychol.*, 1939, **21**, 339-348.
7. LIENERT, G. A. Die Faktorenstruktur der Intelligenz als Funktion des Intelligenzniveaus. Bericht über den 22. Kongress der Deutschen Gesellschaft für Psychologie in Heidelberg, 1959. Göttingen: Hogrefe, 1960.
8. LIENERT, G. A., & EBEL, O. Ein Index zur numerischen Bestimmung der Niveau-Eigenschaft eines psychologischen Tests. *Metrika,* 1960, **3**, 117-127.
9. LORD, F. M. A study of speed factors in tests and academic grades. *Psychometrika,* 1956, **21**, 31-50.
10. THURSTONE, L. L. The Vectors of Mind. Chicago: Univ. Chicago Press, 1944.
11. ———. Multiple Factor Analysis. Chicago: Univ. Chicago Press, 1947.

Psychology Department
The University of Düsseldorf
Düsseldorf, Germany

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 275-286.

A COMPARISON BETWEEN CERTAIN RETINAL AND CORTICAL ACTIVITIES UNDERLYING VISION*[1]

Department of Psychology, Michigan State University

S. HOWARD BARTLEY

A. INTRODUCTION

A sizable fraction of the research carried on within psychological circles in vision concerns itself not only with the phenomenology of vision but also with the body mechanisms involved in producing the phenomena.

One such concern has to do with comparing activities in retina and brain. The matters to be presented here include the following: (*a*) comparisons between electroretinograms and records of cortical response evoked by stimulation of the optic nerve with the eye removed (the stimulation in both cases consisted of a train of photic pulses, otherwise known as intermittent stimulation); (*b*) interpretations of the basis for the differences found in the responses in retina and cortex; and (*c*) a discussion of the expected brightnesses of the members in a series of flashes produced by intermittent stimulation, and the measurement of flash rate in relation to photic pulse rate.

1. *Electroretinograms*

In 1935, Granit and Therman (15) reported upon ERGs (electroretinograms) obtained when using intermittent stimulation. They found that when a train of fairly widely separated pulses was delivered to the eye of the frog, the members of the train were all responded to. At other times, if a rapid train of pulses was delivered, only alternate pulses were responded to. If such a train of pulses was slowed down somewhat, small waves (responses) alternated with the large waves previously obtained. The small waves were then the responses that failed to occur with more rapid photic inputs. As input rate was further slowed down, the small waves progressively increased in size until all the waves were equal.

In 1937, Bartley (3), in studying retinal response to photic pulses in the rabbit, obtained an unexpected result when he plotted pulse interval

---

* Received in the Editorial Office, Provincetown, Massachusetts, on May 2, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported by Public Health Service Research Grant NB 05260-02, from the National Institute of Neurological Diseases and Blindness.

(reciprocal of pulse frequency) against latency. As pulse interval shortened, latency of the waves in record *increased;* but this occurred only up to a certain point and then reversed so that as pulse interval was further shortened, latency *decreased* (see Figure 1).

The explanation given was as follows: Latency is at a minimum with long intervals between stimulus inputs. As intervals are shortened, latency lengthens from encroachment of the inputs on the recovery periods of the responding elements. Soon the elements with the longest recovery times can no longer respond to successive members of the input train. They then will respond only to every second input. Further shortening of the intervals in the input train encroaches further and involves more and more elements. It can cause all elements to respond only to alternate inputs. Latency becomes shorter, again, for such elements because effective input interval is doubled. This then provides for the curve found in Figure 1, and the kinds of records Granit and Therman obtained.

If and when the retina responds unequally to all photic inputs in the

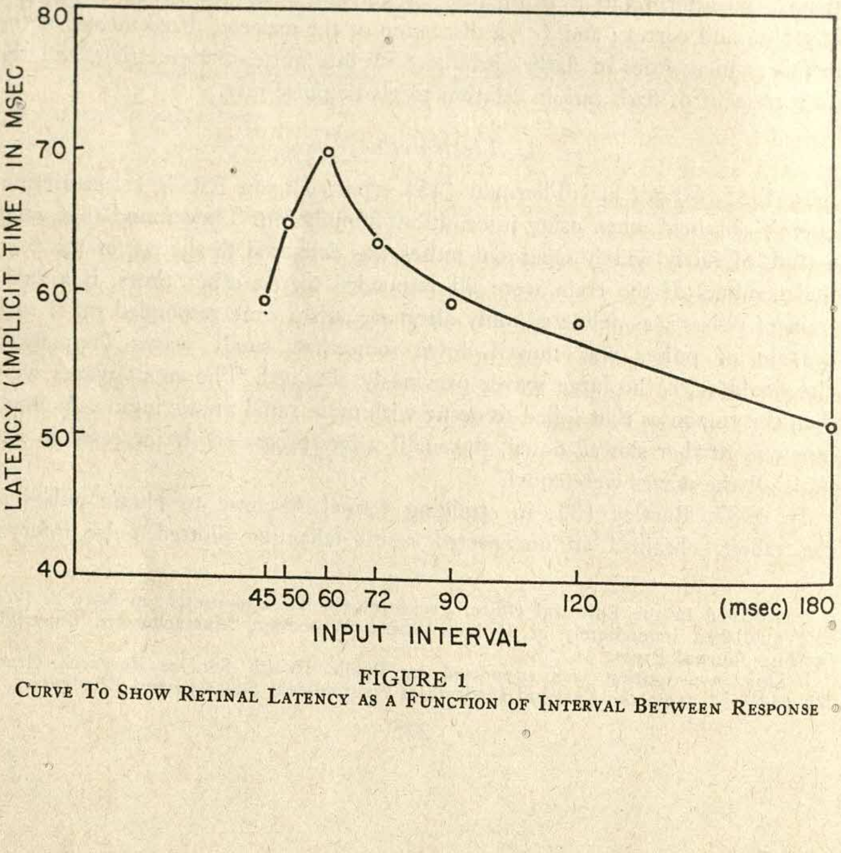


FIGURE 1
CURVE TO SHOW RETINAL LATENCY AS A FUNCTION OF INTERVAL BETWEEN RESPONSE

train, or only to every second photic input, the cortex receives a train of neural inputs that does not tally with the external stimulus. Thus, for part of the frequency range in repetitive stimulation, the cortex probably (even in the human visual system) does not receive a copy of the external stimulus input.

In dealing with the question of how the cortex, or central end of the optic pathway, functions, two lines of consideration are often involved: (*a*) how does the cortex utilize the input it receives? and (*b*) what kinds of inputs does it receive under various external stimulus conditions? These two questions often become confused in discussions of the activity of the optic pathway. Both are involved in the present case, but in the activity model of the cortex to be presented later, it is only the first question that is relevant.

2. *Cortical Responses to Intermittent Stimulation*

An early experiment performed by Bartley (2) best exemplifies what happens when frequencies of intermittent stimulation higher than the intrinsic periodicity of the cortex are delivered to the optic nerve. The rabbit's cortex manifests a periodicity that is often about one-fifth second. Any rate above five per second would be expected to exceed the ability of the elements of the cortex to respond and recover in keeping with the input rate.

The course of events as an input train of, let us say, 18 shocks per second are delivered to the optic nerve, was found to be as follows: The first shock elicited a full-sized response. The second shock elicited none, and from there on until about 12 shocks were delivered, the response was irregular. Some shocks elicited one size of response, some another, and one elicited no response. But, after a period of this irregularity, all shocks elicited responses, and all responses were of the same size though much smaller than the initial response. It was supposed that a reorganization among the responding elements was forced on the system, wherein some elements responded to one shock and other elements responded to others. With complete reorganization, as many elements were responding to any one shock as to its predecessors and successors. The cortex, thus, manifested none of the regular continuous alternation in size of responses as did the records of Granit and Therman's retinal preparation.

B. PROBLEM

The central problem to which we wish to address ourselves concerns the basis for the outcomes in the two segments of the pathway differing in the way just described. In both, there is a form of response redistribution among the various elements involved, but it is different in the two cases. A basis for this is not too difficult to provide.

To make a brief answer before graphically supporting it, we can say that the response difference lies in the fact that the neuro-retina is not repetitive to the same degree as the visual cortex and does not manifest a rhythmical activity-rest cycle such as is characteristic of the cortex. In order to obtain activity in the eye, a stimulus has to be supplied.

Granit (13, 14) has found exceptions to this, but they do not seem to do away with the idea that the elements of the eye are less repetitive and less spontaneously active than the elements of the cortex. The elements of the cortex are called upon to participate in more roles than can be imagined for those in the eye. Some of these even lie outside what is strictly called vision.

Once a brief stimulus has been delivered, the brief discharge in the responding elements of the retina shortly terminates, and there is no activity until another stimulus is delivered. Once the recovery has occurred, subsequent inputs will always find the elements in the same state—ready for new response—rather than being in some phase of repetitive activity.

This is not the case with the cortex. Certain of the elements that are active in constituting the response to specific inputs via the optic nerve are not only activated by this route, but by inputs from other origins. Any input that activates these elements produces a somewhat repetitive response in them. While this repetition may be fairly limited in many cases, there is enough recurrence of input in the cases we are discussing to keep the elements in rhythmic action. When, for any reason, the population is pretty much in synchrony, the alpha rhythm is produced. When specific inputs via the optic nerve are supplied, the extent of repetitiveness and the temporal distribution of the various elements in the population are manipulated.

The response consequences for the two differently acting systems (retina and cortex), as inputs are delivered at various rates, are pictured in the figures that follow.

C. GRAPHICAL MODELS

Figure 2 indicates the consequences for the retina that is supposed to be nonrhythmical: i.e., inactive until and unless excited from the outside as by photic radiation.

The vertical lines bound equal units of time and indicate the successive instants of stimulation. Line 0 is the instant of initial stimulation. There are several horizontal lines, each representing a certain proportion of all of the elements first activated. That is, the bottom line equals 100 per cent; the next line indicates 80 per cent; the next indicates 60 per cent, and so on, with the top line representing zero per cent.

The slanted curved lines represent the recovery times of the various re-

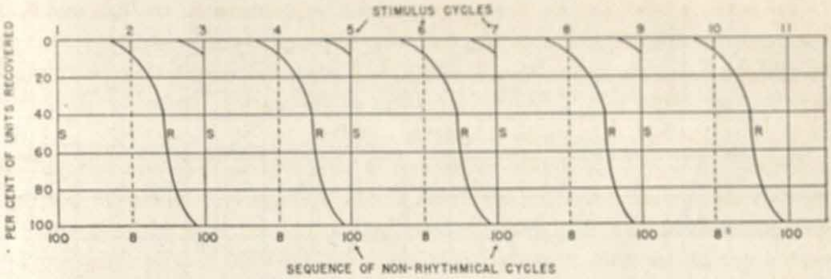


FIGURE 2

A SCHEMA TO SHOW RELATIONS BETWEEN STIMULUS RATE, RECOVERY PERIODS OF RESPONDING UNITS IN THE RETINA (RECEPTOR-BIPOLAR-GANGLION CELL CIRCUITS) AND THE FRACTIONS OF THE UNIT-POPULATION RESPONDING TO SUCCESSIVE STIMULUS INPUTS

The curved lines are the successive recovery cycles of the units. The vertical lines are the instants of stimulation. Horizontal lines represent percentages of total population, from top to bottom. See text for discussion.

sponding units. It will be seen that not all units recover by the time succeeding stimuli (photic pulses) are presented. In the diagram about 8 per cent of the units do recover each time and, thus, can be activated. Hence, for stimulus 2, only 8 per cent of the units respond, for stimulus 3, the same 8 per cent plus the other 92 per cent respond. Stimulus 3 thus activates all units. By arrival time of stimulus 4, only 8 per cent of the units have recovered and so will be activated. Stimulus 5 activates all units. Hence, the sequence of responses is 100 per cent, 8 per cent, 100 per cent, etc.

The same schema will account for why sometimes only alternate inputs are responded to.

Experimental results show that the cortex behaves quite differently. Alternate large and small responses do not appear in cortical records. Figure 3 will theoretically account for this. Actually, a varied array of experimentation has led to the formation of an overall activity description of the central end of the optic pathway which has been called the Alternation of Response Theory (1, 2, 3, 4, 5, 6, 7, 8, 17, 18).

Figure 3 contains the same elements as Figure 2; the vertical lines indicate the points in time at which stimuli are delivered, the shaded area indicates the times covered by the various proportions of the reacting elements, the horizontal lines mark percentages of the elements, and the subsequent curved lines indicate the further repetitions of the cycles of activity of the elements.

The way Figure 3 differs from Figure 2 is in the fact that each activity cycle follows its predecessor "spontaneously." To have a cycle repeat, one does

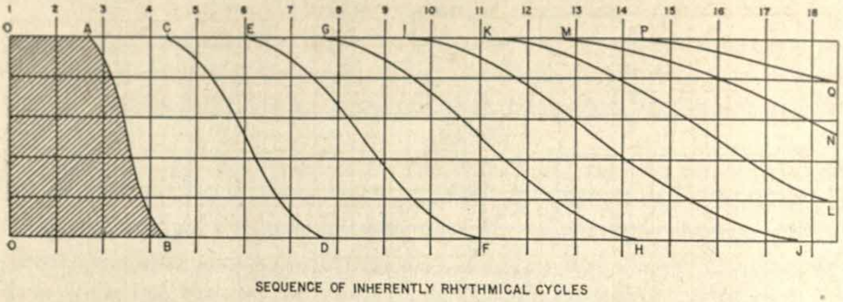


FIGURE 3
Schema To Show Relation of the Same Factors in the Cortex as Indicated for the Retina in Figure 2

not need to deliver a new stimulus from the optic nerve. It is probable that this spontaneous repetitiveness continues for only a few cycles from any single stimulus, but when a train of stimuli is being delivered, the repetitiveness is perpetuated.

It will be noted that as a consequence of the spread in recovery times of the elements and their repetitiveness, the cycles for the element population spread across more and more repetitions of the input (stimulus). For example, while the distance between A-C, C-E, E-G, etc., is constant and represents the elements with the shortest activity-rest cycle, and B-D, D-F, etc., is also constant in representing the elements with the longest activity-rest cycle, the number of repetitions of stimulus input covered at the left of the diagram is two, whereas the number covered by curved line I-J is eight. This means that various stimulus repetitions in this sequence of eight are all finding certain fractions of the population of elements ready to respond.

Figure 4 shows the number of elements responding to each of a number of stimulus cycles. The percentages of elements of all activity-rest cycles are summed at the bottom of the figure for each stimulus cycle. It will be noted that the first input activates the greatest number of elements, which is used as a standard and taken as 100 per cent. The second input activates no elements for none have recovered. The third input activates 10 per cent of the elements. The final cycles activate virtually equal numbers which under the conditions illustrated are about 40-41 per cent. Figure 5 shows this in the form of a graph.

If, for any reason, appreciable numbers of elements were not always to be repetitive, it would prolong the time (extend the number of stimulus cycles) before uniformity in numbers of responding elements would be

ACTIVITY CYCLES OF RESPONDING UNITS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
AB	100	0	0 ↓ 10	10 ↓ 87	87 ↓ 100														
CD					0 ↓ 13	13 ↓ 44	44 ↓ 85	85 ↓ 100											
EF							0 ↓ 13	13 ↓ 35	35 ↓ 63	63 ↓ 89	89 ↓ 100								
GH								0 ↓ 1	1 ↓ 13	13 ↓ 27	27 ↓ 49	49 ↓ 70	70 ↓ 87	87 ↓ 96	96 ↓ 100				
IJ										0 ↓ 1	1 ↓ 10	10 ↓ 25	25 ↓ 41	41 ↓ 59	59 ↓ 74	74 ↓ 87	87 ↓ 95	95 ↓ 100	
KL												0 ↓ 3	3 ↓ 11	11 ↓ 21	21 ↓ 35	35 ↓ 50	50 66		→
MN													0 ↓ 1	1 ↓ 5	5 ↓ 12	12 ↓ 20	20 ↓ 29	29 ↓ 40	→
PQ															0 ↓ 2	2 ↓ 6	6 ↓ 13	13 ↓ 18	→
	100	0	10	77	26	31	54	38	40	41	42	41	42	41	42	44	41	incomplete	

PERCENTAGE OF UNITS RESPONDING

FIGURE 4

THE RESPONSE PERCENTAGES EXPECTED FROM SCHEMA OR MODEL IN FIGURE 3

AB, CD, etc., are the successive activity cycles of the cortical elements. Columns 1, 2, 3, etc., are the successive repetitions of the stimuli. In Figure 5, the values at the bottom of the present figure are depicted as response amplitudes.

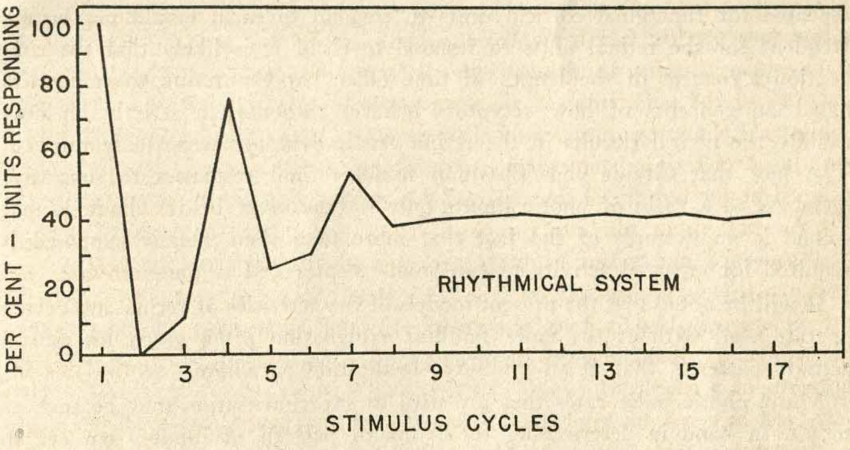


FIGURE 5

A GRAPHICAL REPRESENTATION OF THE RESPONSE AMPLITUDES SHOWN IN THE BOTTOM LINE OF FIGURE 4 responding to each cycle. But the very characteristic of the cortex, as in contrast to the retina, is its far greater tendency to be repetitive. The possibility of this gap of distinction under some natural input conditions becoming narrowed or even temporarily obliterated would do nothing toward invalidating the general distinction being made here.

In viewing the suppositions expressed, it must be kept in mind that the very cortical elements to which we refer may be fired by other sources than the input from the optic nerve. This is certainly the case if it is true that the third (and repetitive) component of the cortical response to optic nerve stimulation involves the same elements as are active in producing the alpha rhythm (10). The supposition that these elements play several roles raises the possibility of certain variations in the overall response picture to optic nerve stimulation. Despite this, cortical response either to electrical stimulation of the optic nerve or photic stimulation of the retina has always been distinct enough and repeatable enough in character in our experiments to make a reasonably firm description of it. This was possible in the day before computers, eliminating chance variables in a series of trials, were invented and used in biological research.

It may have already seemed to the reader that the sorts of comparisons that have been made here between retinal and cortical response are not valid. The retina is supposed to react quite quickly, whereas the cortical cycle involved in the rabbit, for example, is one-fifth second long and in the human, one-tenth second long. Hence, the photic inputs, that might be far too rapidly repeated for functional cortical units to respond to each, would not be too frequent for the retinal units to respond to each. It is likely that the conventional concept of the length of time taken for the retina to respond is based on concepts of how receptors behave, rather than strictly on how rapidly the neural circuits in the retina can individually repeatedly respond. The fact that Granit and Therman found retinal responses to successive members of a train of photic inputs only if they were below about 17 per second is an example of the fact that more time than usually supposed is required for repeated activity to run from receptor cell to ganglion cell.

It will be noted that the present models of the activities of retina and cortex operate well without the more familiar explanation often given for neural activity patterns, that of an inhibition-facilitation paradigm.

Many photic pulse rates that are used in experimentation may be such as to play a hand in determining the temporal pattern of input from eye to cortex. That is to say, many temporal patterns in the grouping of impulses leaving the retina may be quite different in their overall profile from what would be expected if all distributions of impulses were simply a simple reflection of the repetition pattern of the photic input.

The findings and their explanations in terms of the tendency of the retina not to be repetitive and of the cortex to be decidedly repetitive are relevant to two things worth mentioning here: (*a*) The alternation of response—i.e., temporal redistribution of the activity of the members of a response population—is a phenomenon applying both to retina and to cortex. In fact, the same sort of interpretation of findings in other neural preparations has been made. Echlin and Fessard (12) in their description of the way the kinesthetic system functions found that the high discharge rates in the response record were not representative of the rates at which single elements could respond, but rather represented a composite of elements discharging at lower rates. Rose and Mountcastle (22) described the relation between the ability of the tactile system as a total and the ability of its constituent elements to respond. This description resembles, in essence, the description we have given of alternating elements. Moreover, the familiar volley theory of Wever and Bray (26) is an early example (in audition) of the same thing. (*b*) The findings dealt with in the models here ought to bear on the nature of observations made by the human subject when confronted with intermittent photic stimulation.

D. SENSORY END RESULTS

The following seem to be some of the relevant findings and conclusions regarding human visual response. The results we have reference to here are those observations that the human subject makes when the rates of repetition of photic pulses vary, let us say, from one or two per second up to halfway to CFF (critical flicker frequency). Two features of what is seen are of relevance. The one is the brightness of the flashes and the other is the perceived flash rate. Most experiments on these matters do not involve observations produced by the first few photic pulses delivered, but rather the results after the visual mechanism of the subject has settled down to responding in some steady manner.

Nevertheless, some studies have been concerned with various features of response to the very first few pulses (9, 20, 27). The results of these may be summarized for present purposes as indicating that visual response to the first few pulses in a sequence is different from visual response to subsequent pulses. The measure that could be called critical flicker interval (the reciprocal of CFF) is increased, and the first few pulses are responded to in an irregular fashion. This tallies with the description given earlier in this paper for the cortical response findings when the optic nerve was stimulated. Thus, a parallelism between cortical response and sensory (visual) response in the respects referred to is shown.

To what extent is the observer able to report upon such a matter as brightness when presented with a train of photic pulses? Might not the characteristics of what is seen be too subtle and uncertain to be reliably reported upon? The uncertainties may pertain both to brightness and to flash rate, for the two are quite closely associated in appearance.

While, in general, long trains of photic pulses of equal luminance produce sequences of apparently uniform flashes—with no noticeable discrepancies between flash rate and pulse rate as far as the observer is concerned—the author has, at times, had the impression of seeing flashes at half the pulse rate. This is plausible from what has already been described about retinal response. Even if half-rate flashes were to occur, it would not nullify anything said about the close relation between the nature of cortical response and the nature of the sensory end result, for visual response at half rate would be the result, for the cortex itself, of receiving inputs at half rate while responding to every input.

Efforts of various sorts have been made by several authors to measure perceived flash rate relative to photic pulse rate (11, 19, 21, 24, 25). The possibility that has just been pointed out may have some bearing on the findings of these authors.

All of the authors ended up by reporting certain kinds of discrepancies between photic pulse rate and perceived flash rate. All of the studies, except possibly one, likewise found reason to doubt the validity of the measures made in actually representing true flash rates.

The safest conclusion to be drawn when all of the studies are taken into consideration is that after input rates rise about three or four per second, the human observer cannot be relied upon to count numbers of flashes in a limited train or to make measures of their rate by use of some standard such as matching flashes to auditory clicks. While the matching may be possible in some cases, the same fault exists with observing clicks. The deficiency in dealing with clicks in a temporal fashion may not be as great as with flashes but, nevertheless, the deficiency exists and invalidates the method as an absolute reference by which to deduce flash rate.

In the midst of the attempt of the observer to match flashes with clicks, curious phenomenological effects occur (16). These will not be discussed, but it is well-known that the impressions of brightness shift at various ranges in the input rate. Whereas the flashes may stand out from the dark intervals between when rates are low, they begin to lose some of their identity as rate rises, or as the durations of the photic pulses change, or as the PCFs (pulse-to-cycle fractions) change. Hence, the observer is forced to change criteria as a consequence. One study that is relevant here (23) makes this quite certain.

E. Summary and Conclusions

The foregoing studies have served pretty well to isolate and compare one of the essential differences between retinal and cortical activity. While both systems are flexible enough to manifest quite varied reactions or activities, they both display certain characteristics which when examined serve to explain the kinds of results obtained in neurophysiological records and visual observation itself.

The present paper has demonstrated some of these characteristics. In general, it appears that the elements of the retina under undisturbed conditions are not continuously active and inclined toward repetitiveness, whereas certain elements of the cortex involved in the response to inputs via the optic nerve are repetitive. This very difference seems to go a long way in accounting for the differences in the records obtained from retina and cortex.

The discussion of these results has included some indication of what might be expected of visual observation under similar input conditions. These expectations and the findings seem to agree pretty well.

References

1. Bartley, S. H. A comparison of the electrogram of the optic cortex with that of the retina. *Amer. J. Physiol.*, 1936, **117**, 338-348.
2. ———. Temporal and spatial summation of extrinsic impulses with the intrinsic activity of the cortex. *J. Cell. & Comp. Physiol.*, 1936, **8**, 41-62.
3. ———. Some observations on the organization of the retinal response. *Amer. J. Physiol.*, 1937, **120**, 184-189.
4. ———. Subjective brightness in relation to flash rate and the light-dark ratio. *J. Exper. Psychol.*, 1938, **23**, 313-319.
5. ———. Visual sensation and its dependence on the neurophysiology of the optic pathway. *Biol. Symp.*, 1942, **7**, 87-106.
6. ———. Some facts and concepts regarding the neurophysiology of the optic pathway. Symposium on Visual Mechanisms, Bethesda, Maryland, September, 1957. *A.M.A. Arch. Ophthalmol.*, 1958 (Part II), **60**, 775-791.
7. ———. A clarification of some of the procedures and concepts involved in dealing with the optic pathway. In R. Jung and H. Kornhuber (Eds.), *The Visual System: Neurophysiology and Psychophysics*. Symposium, Freiberg, Germany. Berlin: Springer-Verlag, 1961.
8. Bartley, S. H., & Nelson, T. M. Some relations between end results and neural activity in the optic pathway. *J. of Psychol.*, 1963, **55**, 121-143.
9. Bartley, S. H., Nelson, T. M., & Ranney, J. E. The sensory parallel of the reorganization period in the cortical response in intermittent retinal stimulation. *J. of Psychol.*, 1961, **52**, 137-147.

10. Bartley, S. H., O'Leary, J., & Bishop, G. H. Differentiation by strychnine of the visual from the integrating mechanisms of the optic cortex in the rabbit. *Amer. J. Physiol.*, 1937, **120**, 604-618.
11. Cheatham, P. G., & White, C. T. Temporal numerosity: I. Perceived number as a function of flash number and rate. *J. Exper. Psychol.*, 1952, **44**, 447-451.
12. Echlin, F., & Fessard, A. Synchronized impulse discharges from receptors in the deep tissues in response to a vibrating stimulus. *J. Physiol.*, 1938, **93**, 312-334.
13. Granit, R. Rotation of activity and spontaneous rhythms in the retina. *Acta. Physiol. Scand.*, 1941, **1**, 370-379.
14. ———. Sensory Mechanisms of the Retina (reprinted). New York: Hafner, 1963.
15. Granit, R., & Therman, P. O. Excitation and inhibition in the retina and in the optic nerve. *J. Physiol.*, 1935, **83**, 359-381.
16. Knight, R. R. A methodological investigation of perceived flicker rate. Master's dissertation, Department of Psychology, Michigan State University, East Lansing, Michigan, 1966.
17. Nelson, T. M., & Bartley, S. H. Description of a central mechanism related to diverse brightness phenomena. *J. of Psychol.*, 1964, **58**, 379-395.
18. ———. Theoretical interpretation of various qualitative and quantitative aspects of flicker and fusion phenomena. *J. of Psychol.*, 1964, **59**, 185-194.
19. Nelson, T. M., Bartley, S. H., & Bochniak, F. Flash frequency when brightness is varied by manipulation of temporal distribution of photic intermittency. *J. of Psychol.*, 1964, **59**, 195-205.
20. Nelson, T. M., Bartley, S. H., & Jewell, R. M. Brightness changes associated with variation in the number of pulses in a repetitive photic train: Sensory implications of the Alternation of Response Theory, II. *J. of Psychol.*, 1963, **56**, 107-113.
21. Reese, T. W. The application of the theory of physical measurement to psychological magnitudes with three experimental examples. *Psychol. Monog.*, 1943, **55**, 1-89.
22. Rose, J. E., & Mountcastle, V. B. Touch and kinesthesis. In H. W. Magoun (Ed.), *Handbook of Physiology* (Sec. 1, Vol. I). Washington, D.C.: Amer. Physiol. Soc., 1959.
23. Schneider, C. W., & Bartley, S. H. Changes in sensory phenomena and observer criteria at low rates of intermittent photic stimulation. *J. of Psychol.*, 1966, **63**, 53-66.
24. Shipley, T. Auditory flutter-driving of visual flicker. *Science*, 1964, **145**, 1328-1330.
25. Taubman, R. E. Studies in judged number: II. The judgment of visual number. *J. Gen. Psychol.*, 1950, **43**, 195-219.
26. Wever, E. G., & Bray, C. W. Present possibilities for auditory theory. *Psychol. Rev.*, 1930, **37**, 365-380.
27. Wilkinson, F. R. The organization of the visual response. Doctoral dissertation, Department of Psychology, Michigan State University, East Lansing, Michigan, 1961.

Department of Psychology
Michigan State University
East Lansing, Michigan

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 287-290.

NEUROPSYCHIATRIC PATIENTS, ACCIDENT PRONENESS, AND INTEREST PATTERNS*[1]

University of Washington and Veterans Administration Hospital,
Jefferson Barracks, Missouri

JOSEPH T. KUNCE AND BLAYNE BREWER

A. PROBLEM

The incurrence of accidents has often been blamed upon an agitated or upset state of mind. Although this may be true for people under situational stress, it is the authors' opinion that accidents generally reflect a style of life and, therefore, "accident proneness" can be predicted. Certain characterological facets of behavior may be of more relevance to accident proneness than are chronic mental or emotional disturbances. In fact, individuals with a tendency to worry excessively may have substantially fewer accidents than people in general.

In a preliminary study on 84 physically disabled subjects, a significant relationship between an individual's measured interest patterns and the *severity* of his incurred disability was found [Kunce and Worley (3)]. Cross-validation studies showed a significant relationship of these interest patterns to the *frequency* of accidents at school over a 3-year interval for 146 high school students prior to testing [Kunce *et al.* (5)], and over a 2-year interval for 112 University students subsequent to testing [Kunce and Worley (4)]. This research was designed to investigate the relationship of interests with accident behavior among a neuropsychiatric population.

B. METHOD

1. *Subjects*

Records of 189 chronic neuropsychiatric male patients at the Veterans Administration Hospital, Jefferson Barracks, Missouri, who had taken the Strong Vocational Interest Blank (SVIB) either for counseling or research purposes, were used to provide a source of subjects.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on May 16, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This project was supported in part by Vocational Rehabilitation Administration Grant No. RT-3 and Public Health Service Grant, #CH 51-02-B-64.

2. *Instruments*

A psychological index of accident proneness (AP) was obtained by subtracting the SVIB standard scores on the banker scale from the aviator scale. This method was also used in the aforementioned studies. These particular scales were chosen on the hypothesis that high aviator scores would be related to adventurousness and the likelihood of accidents, and high banker scores to cautiousness and avoidance of accidents.

A list of the names of these subjects was given to members of the psychology staff ($N = 7$) to rate patient accident behavior. All of these judges had been members of the hospital staff for at least three years. They were asked to indicate (for each patient they knew) the following: Yes (Y) if the subject did, or No (N) if he did not, have an accident or injury during his course of hospitalization. A total of 60 subjects were rated by more than one staff member, and 56 by at least one member of the staff, giving a total of 116 subjects. We used this method for getting a measure of accident proneness because it was not possible to determine accurately from hospital records either the frequency or severity of all the accidents that any one subject may have had in a given time interval. In a sense, the procedure used gives an index of whether the subjects were *perceived* as accident prone or not.

3. *Procedure*

The relationship of AP scores with perceived accident behavior was evaluated separately for the group of subjects that received single ratings and for those with mutiple ratings. For each group, subjects were categorized according to AP scores (+11 up, +10 to —10, or —11 down) and previous accidents (Y only or $Y > N$ and N only or $N \geqslant Y$). Chi square was computed upon the distribution of frequencies and one-tailed test for significance applied, since it was predicted upon basis of previous research that high AP scores would be positively related to incurrence of accidents.

C. RESULTS

The two-by-three distributions for the group having single judgements yielded a chi square of 5.20, and for the multiple judgement group they yielded a chi square of 4.21 (see Table 1). The chi square for the combined groups equaled 9.45, which is significant at the .01 level of confidence. In the group of subjects having AP scores of 11 up, 56 per cent were perceived as accident prone, in comparison to 29 per cent in the intermediate group and 13 per cent in the group with the lowest AP scores.

TABLE 1
SUBJECTS CATEGORIZED BY ACCIDENT-PRONENESS SCORES AND HISTORY OF ACCIDENTS

Group	AP scores	History of accidents Yes	No	Total	Chi square	Sig.
Single ratings ($N = 56$)	11 up	2	0	2	5.20	.05
	±10	11	27	38		
	—11 down	3	13	16		
Multiple ratings ($N = 60$)	11 up	7	7	14	4.21	.05
	±10	9	23	32		
	—11 down	1	13	14		
Total ($N = 116$)	11 up	9	7	16	9.45	.01
	±10	20	50	70		
	—11 down	4	26	30		

D. DISCUSSION

These results are consistent with our previous research showing a significant relationship between interest patterns and various criteria of accident proneness. However, in the previously mentioned studies 48 per cent of physically disabled subjects, 47 per cent of the high school students, and 56 per cent of the university students had AP scores of 11 up in contrast to 14 per cent of the neuropsychiatric subjects. Superficially, these results suggest that chronic and severe maladjustment and accident proneness may be incompatible.

Such a conclusion conflicts with those of Selzer *et al.* (6), who reported a high incidence of recognized psychiatric illness among accident-prone drivers. In their study, the majority of subjects diagnosed as "mentally ill" were classified as having characterological disorders and less than 5 per cent as schizophrenic. Perhaps too broad (or narrow) a definition of mental illness is responsible for the conflicting findings. Reference to two other studies helps clarify the possible role of personality characteristics. Research done by Fordyce (2) demonstrates a relationship of characterological traits and imprudent behavior prior to the incurrence of severe physical injuries, and Fine (1) reports a significant relationship to extraversion and incurrence of accidents. All of these data support a viewpoint that various outgoing personality traits are related to incurrence of accidents and, conversely, more conservative and withdrawing behavior to avoidance of accidents.

We also re-examined the data on the university group for clues of the relationship of accidents to emotional distress. We found that 10 of the 112 students had been referred for psychiatric care. The accident rate of the 10 subjects was identical to that of the other students, and of these 10 students only those who had high AP scores had recorded accidents during a two-year interval following testing.

Since we believed that "adventurousness" could be a factor in the incurrence of accidents, we also investigated the relationship of the AP scores with the chance of being known by the psychology staff. Of the 16 subjects having high AP scores, 88 per cent were known to more than one judge in contrast to the 46 per cent of those with lower scores. This difference is significant at the .01 level.

In conclusion, these findings are consistent with the hypothesis that chronic emotional disturbance may not be as significant a factor in accident proneness as other unique aspects of personality. Further research to clarify the nature of accident proneness is needed to determine whether AP scores simply reflect a person's overall activity level or more unique psychological traits.

REFERENCES

1. FINE, B. J. Introversion-extraversion and motor vehicle driver behavior. *Percept. & Motor Skills,* 1963, **16**, 95-100.
2. FORDYCE, W. E. Personality characteristics in men with spinal cord injury as related to manner of onset of disability. *Arch. Phys. Med. & Rehab.,* 1964, **45**, 321-325.
3. KUNCE, J. T., & WORLEY, B. H. Interest patterns and disability. *J. Clin. Psychol.,* 1966, **22**, 105-107.
4. ———. Interest scores and accidents among university students. *J. Coll. Student Person.,* in press.
5. KUNCE, J. T., STURMAN, J., LONGHOFER, P., & CASTOR, M. Psychological factors in high school accidents. *Person. Guid. J.,* in press.
6. SELZER, M., PAYNE, C., GIFFORD, J., & KELLY, W. Alcoholism, mental illness and the drunk driver. *Amer. J. Psychiat.,* 1963, **120**, 326-331.

School of Medicine
Department of Physical Medicine and Rehabilitation
University of Washington
Seattle, Washington 98105

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 291-297.

A METHOD FOR STUDYING OLFACTION IN THE UNRESTRAINED RAT*[1]

Department of Psychology, Tufts University

WARREN H. TEICHNER

A. INTRODUCTION

When an animal is presented with an odorant, it can be assumed that it first *detects* the odorant and that it then *evaluates* it in some way. While both responses may be assumed to depend upon the concentration of the odorant in air and on the state of receptor adaptation, evaluation is a response to suprathreshold concentrations that depends in addition on the attractiveness or aversiveness of the compound. Presently available methods for studying olfaction in animals depend on either electrophysiological techniques or on conditioned responses using the odorant as a conditioned stimulus. Both methods are limited to inferences about detection, although if a behavioral measure of evaluation were available, meaningful electrophysiological correlates of it might be obtained. The present report is of a new method for the measurement of olfactory evaluation; the method provides for the measurement of detection as well.

It is easy to observe that a characteristic response of the rodent to any alteration of its environment, including the presentation of an odorant, is sniffing. Sniffing is clearly distinct from breathing; it is audible and it is visible in terms of twitching of the animal's nose and whiskers. Whereas odorants may be detected by the animal during breathing, it is while sniffing that they appear to be evaluated. Furthermore, the latency of the initiation of sniffing following presentation of an odorant presumably reflects the speed of detection during ordinary respiration. Sniffing, then, is suggested as

---

* Received in the Editorial Office, Provincetown, Massachusetts, on May 17, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This paper is a report of research undertaken in cooperation with the U. S. Army Natick Laboratories under Contract DA 19-129-AMC-236N. The findings in this report are not to be construed as proof of Department of the Army position. The encouragement of Dr. John Pratt is acknowledged. The PNP, impregnated on burlap, was supplied by the U. S. Fish and Wildlife Service, Denver, Colorado. The author is also indebted for the technical assistance of Joseph Chereski in the early stages of apparatus development and to the later technical assistance of Ann Gustin, Theodore Nalwalk, and Mrs. Leah Price.

a means for describing the animal's olfactory detection and evaluation. A first step toward the development of a sniff methodology will be described.

Although it would be possible to record both sniffing and breathing by available techniques, all such techniques available impose some form of restraint to respiration itself. Furthermore, restraint acts as an emotional stimulus and it is likely that sniffing is one of the characteristic responses of the rodent to emotional stimuli. It was considered desirable, therefore, to develop a method for measuring sniffing that imposes no restraint on the animal. The simplest way to do this seems to be by audio-recording of sniff sounds. This was the approach taken.

B. Hypotheses

A number of simple hypotheses can be generated about the relationship between sniffing and the attracting and repelling characteristics of odorants. As listed below, these hypotheses assume some baseline level of ongoing sniffing prior to presentation with the odorant:

1. The rate of sniffing and the amplitude or intensity per sniff should decrease with stimulation by repellents and increase with stimulation by attractants. The amount of change should be a function of the degree of repellency or attractiveness of the odorant.

2. With continued constant stimulation, sniffing should adapt; that is, the rate and amplitude of sniffing should return to the baseline level. The rate of adaptation should be a function of the attractiveness or repellency of the odorant.

3. Removal of the odorant provides a new stimulus condition and, therefore, should affect sniffing. Assuming complete adaptation, since removal provides a nonrepellent stimulus, removal should be followed by an increase in sniffing rate and amplitude regardless of the nature of the previously presented odorant. However, following removal of an aversive stimulus, sniff rate and amplitude should be greater and adaptation should be slower than following removal of an attractant.

4. Regardless of whether the stimulus is an attractant or repellent, sniff rate and amplitude should be inversely related to the concentration of the odorant in the air. This is based on the assumption that, for an air sampling mechanism, the weaker the concentration, the greater must be the air sample required for a decision.

5. The speed of occurrence of the first sniff following presentation or removal of an odorant should be greater for repellents than for attractants.

6. The speed of response of the first sniff following presentation or re-

moval of the odorant should be directly related to the concentration of the odorant in the air.

The above hypotheses are based partly on intuition and partly upon the general nature of sensory phenomena. They are all testable, but their goodness is less immediately important than their utility as working hypotheses. The present report provides some data bearing upon those hypotheses as they concern repellents. However, our primary purpose in this report is to provide a first test of sniffing as a technique.

C. APPARATUS

The apparatus used is shown schematically in Figure 1. The animal was placed in a glass enclosure with a 8.5 cm diameter and height of 23 cm, inside dimensions. Ascending vertically, the neck of the enclosure began at 16.5 cm from the bottom and sloped to center to form an outwardly lipped orifice. The inside diameter of the orifice, excluding the lip, was 1 cm. A ball joint with an inside duct equal in size to the orifice was placed on the lip of the

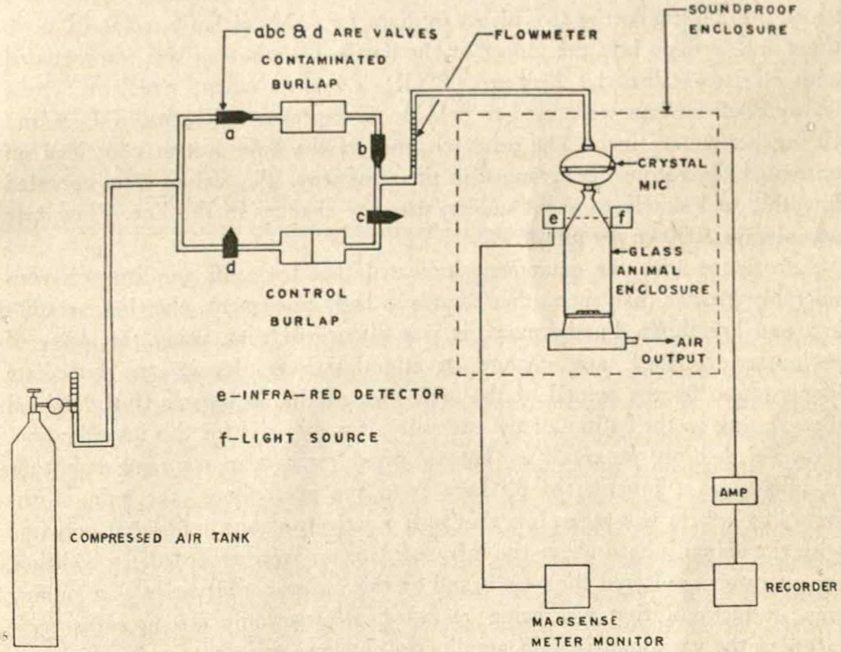


FIGURE 1
SCHEMATIC OF THE APPARATUS

neck and connected to a second glass enclosure above. This design permitted easy removal of the animal enclosure. The second enclosure contained a crystal microphone, the back of which baffled the air arriving from the flow system shown in the figure. Both glass enclosures were housed in a highly sound-dampened box containing a window through which the enclosed animal could be observed. An infrared detector provided a recorded indication of the position of the animal, whether above or below the placement indicated. Sounds picked up by the microphone were amplified and recorded on a Beckman-Offner-Dynagraph at a paper speed of 2.5 cm per second.

D. Procedure

The animal was inserted through the open bottom of the animal enclosure, which was then closed by a rubber stopper. Embedded in the stopper was a removable plastic grid on which the animal rested. Use of the grid allowed for an air space below the animal, which prevented it from blocking the outgoing air flow.

The flow system contained two channels. In each channel, burlap was sealed into plastic bottles two inches in diameter. The burlap was folded three times and wedged into the center of the bottle. One burlap was impregnated with 2-Nitro-1 Phenyl-1 Propene (PNP), a known rodent repellent. Three sealed PNP burlaps were used to provide an experimental series of 1, 3, and 10 mg per square inch. The other channel of the flow system contained an untreated control burlap arranged in the same way. The valves were operated smoothly and slowly to avoid sudden pressure changes in the line. Flow rate was always 2000 cc per minute.

Experience with the equipment indicated that the sniff amplitude is considerably greater than such other sounds as body movement, chewing, scratching, and breathing. Furthermore, it was discovered that, using the shape of enclosure described, any change in stimulation results in an immediate investigation by the animal of the neck area of the enclosure. If the animal were resting on the bottom of the enclosure, it would rise for this investigation. However, because we could not control for differences in recorded amplitude of sniffing as a function of distance from the microphone, the gain of the recording system was set so low that only relatively strong sniffs that occurred with the animal's head above the infrared detector were recorded. In addition, the operator monitored the sounds and by this means could verify that sniffing was, in fact, all that was being recorded. Because only strong sniffs were recorded the variations in sniff amplitudes obtained were not analysed.

Prior to the experiment proper, the animals were habituated to the apparatus. This consisted of 10 days of enclosure for five minutes per day with the control burlap. Following this each animal was tested in a continuous pre-exposure-exposure-postexposure series. When the animal appeared to be below the recording position and resting, the operator began a 15-second control burlap period (pre-exposure), then switched to a 15-second PNP period (exposure), then switched back to a final 15-second control burlap period (postexposure). When making these changes, the operator pressed a button that controlled an event marker on the recorder. The three PNP levels were presented to all of the animals in a Latin square sequence. The glass animal enclosure and the sound-dampened box were aired out for a minimum of five minutes between individual runs. This was aided by a strong fan. The enclosure was washed thoroughly following each daily session.

Preliminary investigation suggested that an important amount of adaptation might result from a single test series and that the adaptation might persist for several hours. As a result each animal was tested only once in a single day. Earlier experimentation also suggested that the rate of sniffing is greater for hungry animals than for animals that are not hungry. Therefore, the animals were placed on a 23-hour food deprivation schedule starting three weeks prior to exposure to the PNP. Testing was always at the same portion of the day and within two hours of feeding.

Twelve male albino rats of the CD strain were used as subjects. The animals were five months old at the start of testing. They had had some prior experience with the apparatus and with a variety of odorants in preliminary work. Their use for this purpose was based on the observation that there seemed to be no persistence of sniff adaptation to the various preliminary odorants used.

E. RESULTS

The data were analyzed in terms of the number of sniffs per second and the latency of the first sniff of the exposure period and of the postexposure period. On this basis the data showed important individual differences in sniffing.

In order to assess the results, minimizing individual differences, the number of sniffs produced per rat in the last five seconds of the pre-exposure period (A), the number of sniffs in the first five seconds of the exposure period (B), and the number of sniffs in the first five seconds of the postexposure period (C) were counted. On this basis the percentage change of B from A and of C from A was calculated. The results are shown in Figure 2 as a function of concentration. In this figure values above zero represent an increase in sniffing relative to the end of the pre-exposure period, whereas values below zero represent a decrease. As the figure shows, number of sniffs was inversely related to concentration in both the exposure and postexposure periods. The figure also shows that the relative increase was greater for the postexposure period than for the exposure period, and that the range of percentage change was very large.

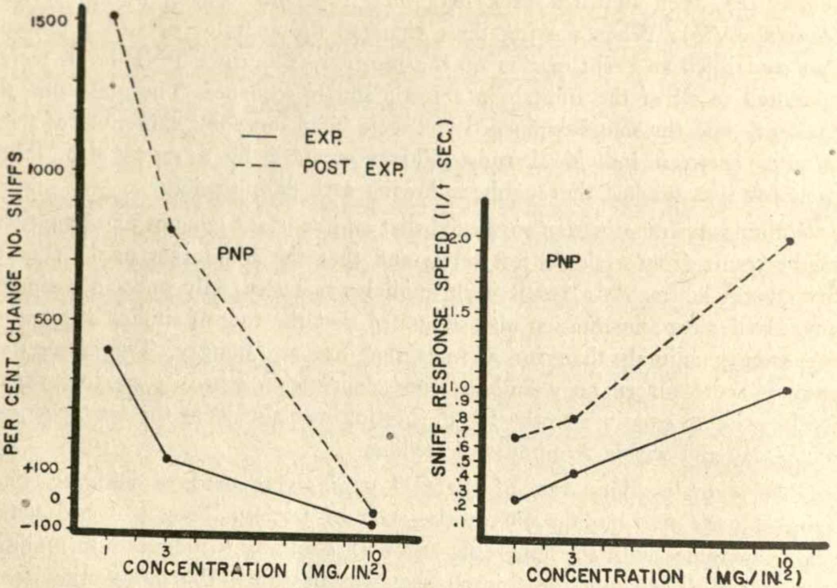


FIGURE 2
MEAN PER CENT OF CHANGE IN NUMBER OF SNIFFS DURING THE FIRST FIVE SECONDS RELATIVE TO THE END OF THE PRE-EXPOSURE PERIOD (LEFT) AND MEAN RECIPROCAL SNIFF RESPONSE TIME (RIGHT)

Each is shown as a function of concentration for the exposure and postexposure periods.

Figure 2 also presents the speed with which the first sniff occurred in the exposure and in the postexposure periods. The data are the means of the reciprocal of the latency of the first sniff and, thus, reflect response speed rather than time. The figure shows that the speed of occurrence of the first sniff was directly related to the concentration and that it was greater in the postexposure period than in the exposure period.

F. DISCUSSION

As noted above, the data are based only on strong sniffs that occurred near the microphone. We are presently trying to develop an improved version of the apparatus that, in addition, will also permit analysis of sniff amplitude. However, the present data are sufficiently clear to indicate the potentialities of the method and to suggest that, at least as concerns both detection (speed of first sniff) and suprathreshold evaluation (number of sniffs) it offers a method for quantifying olfactory repellency. In this regard, the data suggest, for PNP, that this compound is not repellent until some concentration between 3 mg per square inch and 10 mg per square inch. The results also suggest, as hypothesized, that the number of sniffs is inversely related to concentration of repellent and that the speed of sniffing is directly related.

In addition to its possibilities as a method for studying olfaction in rodents, and possibly other animals, sniffing may provide a useful means for studying the orienting response and its habituation, exploratory behavior, intersensory stimulation, and the effects of combined primary drive conditions upon a single response. That is, with regard to the latter, if, as we observed, hunger affects sniffing, there is the possibility of studying the combined effect of hunger and thirst or other such conditions on sniffing. Finally, the method appears expedient in that it requires only habituation of the animal to the enclosure rather than extensive and difficult behavioral conditioning procedures.

Published as a separate and in *The Journal of Psychology*, 1966, **63**, 299-308.

PERCEIVED DISCREPANCY IN ATTITUDE INTENSITY BETWEEN THE ACTORS A AND B IN ABX SITUATIONS AND ITS EFFECT UPON AFFECT

OR

WHY DON'T EXPERIMENTERS UTILIZE ALL THEIR AVAILABLE DATA IN TESTING HYPOTHESES?* [1]

Institute for Defense Analyses, Arlington, Virginia

NEHEMIAH JORDAN

A. INTRODUCTION

A recent paper [Price, Harburg, and McLeod (6)] reports on an investigation of the effects of changes in attitude intensities in ABX situations[2] upon how such situations are experienced by subjects in terms of "pleasantness" and "unpleasantness." The ABX system was formulated by Newcomb and discussed rather extensively in at least three publications (2, 3, 4). In these publications, the discussion devolves exclusively upon the "strain" experienced by an actor, A, when he perceives a discrepancy to exist between the signs of his attitude towards some entity or person, X, and the attitude of another person, B, with whom he interacts. This "strain" is experienced as "unpleasantness." Little if anything is said about how ABX situations are experienced when they are not "in strain." Following the more general contemporary conceptualizations of cognitive balance,[3] the investigators assume that an ABX situation "not in strain" is balanced and will therefore be experienced as being "pleasant."

The important novelty of the present research lies in its attempt to investigate the effect of changes in attitude intensity. Research on such triadic groupings has generally concentrated on global attitudes: e.g., A's general attitude towards B (or the mutual attitudes between A and B), A's general attitude towards X, and B's general attitude towards X. However, attitudes

---

* Received in the Editorial Office, Provincetown, Massachusetts, on May 18, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] Previously published as Internal Note N-282(R), Institute for Defense Analyses, Economic and Political Studies Division, Arlington, Virginia, November, 1965.

[2] An ABX situation is a social situation comprising an actor, A; another person, B; and an entity, X (which may be a third person). It is defined in terms of the mutual attitudes between A and B, and their respective attitudes towards X.

[3] For a good review of these contemporary conceptualizations, see Brown (1, Ch. 11).

vary in intensity; yet the effect of changes in intensity of attitudes on the experienced pleasantness or unpleasantness of such groupings has all but been neglected in contemporary research. Granting that balanced configurations are experienced as being "pleasant," then, all other things being equal, the stronger the attitudes of the actors in such situations, the more "pleasant" these situations should be experienced; granting that imbalanced configurations are experienced as being "unpleasant," again, all other things being equal, the weaker the attitudes of the actors involved, the less "unpleasant" the situation should be experienced.

But all other things are not equal. As long as contemporary research restricted itself to global attitudes—A's liking ($+L$) or disliking ($-L$) of B, etc.,—only one source of discrepancy leading to "strain" or imbalance could be studied: the differences in attitude signs. With the introduction of attitude intensity a second source of discrepancy arises: although both A and B may share the same attitude towards X in terms of signs, the intensity of their attitude may vary—e.g., A may like X very much while B merely likes X moderately. Differences in attitude intensities can also be a source of conflict, of imbalance, or of "strain." Their effect can be conceptualized as "weakening" the general configuration, thereby also causing "basically pleasant" configurations to be experienced as "less pleasant" and basically "unpleasant" configurations to be experienced as "less unpleasant."

B. HYPOTHESES

These considerations lead to the formulation of two general hypotheses with respect to changes in the experienced affect of ABX situations resulting from changes in attitude intensity: (*a*) Keeping the attitudes of A and B towards X constant, the degree of experienced pleasantness of basically pleasant balanced ABX situations and the degree of experienced unpleasantness of basically unpleasant imbalanced ABX situations are a direct function of the intensity of attitude ($+L$) between A and B. (*b*) Keeping the attitude intensity ($+L$) between A and B constant, the degree of experienced pleasantness of basically pleasant balanced ABX situations and the degree of experienced unpleasantness of basically unpleasant imbalanced ABX situations are an inverse function of the degree of discrepancy between the attitude intensities of A and B towards X.

In the paper under discussion, these two hypotheses are formulated thus:

> 1. Given a constant degree of perceived discrepancy, positive or negative affect varies *with the degree* of attraction between A and B.
>
> 2. Given a constant positive A to B attraction, the amount of positive

> or negative affect varies *with the degree* of A's perceived discrepancy (6, p. 90).[4]

The general technique used to test these hypotheses was to have subjects rate imagined AB*X* triads on a graphic rating scale, a technique common in this research area. The novel independent variable consisted of three attitude intensities: like strongly, $+++L$; like mildly, $++L$; and like slightly, $+L$; and conversely, dislike strongly, $---L$; dislike mildly, $--L$; and dislike slightly, $-L$. Unfortunately, rather arbitrarily, not all possible combinations of these variables were presented to subjects for rating. The investigators write:

> Logically, the number of ABX situations where the sign and/or intensity of relations among three persons are systematically varied is high, but may be limited by the range of intensities of attitudes involved. We actually used forty-four situations because of time limitations (6, p. 91).[5]

In testing the predictions of these hypotheses, the investigators compared the ratings of situations with strong intensities: like strongly ($+++L$) or dislike strongly ($---L$), with the ratings of the comparable situations with the weakest intensities: like slightly ($+L$) or dislike slightly ($-L$), in 51 of 60 paired comparisons; like strongly or dislike strongly was compared to like mildly ($++L$) or dislike mildly ($--L$) in five paired comparisons; and like or dislike mildly was compared to like or dislike slightly in four paired comparisons. The investigators report that 70 per cent of the paired comparisons supported the predictions at the .05 level or better. The hypotheses appear to be tenable.

But are they? The hypotheses, as formulated by the investigators, explicitly state that the effects of changes in attitude intensity vary *with the degree* of the changes—i.e., the greater the attitude change, the greater the effect. Testing these hypotheses by means of paired comparisons—i.e., only two values or one change—is a weak test. Were more values available the test would be more stringent. Since the subjects were presented with three degrees of attitude strengths—strong, mild, and slight—more values are available. The reported data can be rearranged to provide 43 sets of situations where everything is kept constant except for a change of attitude intensity from slight, through mild, to strong. This provides three values instead of two, and enables a more stringent test of the hypotheses by using all the data. The results of doing this are shown in Tables 1 and 2.

---

[4] Italics inserted by the present author. These two general hypotheses are then expanded to seven more specific hypotheses to fit the special AB*X* situations to be tested. The expansion is not of importance to the present discussion.

[5] Fifty-eight situations were actually used. No explanation is given for this discrepancy.

TABLE 1

THE EFFECTS OF VARIATION OF ATTITUDE INTENSITY ON ABX SITUATIONS OF THE FORM: A HAS A POSITIVE ATTITUDE TO B; A HAS A POSITIVE ATTITUDE TO X; B HAS A POSITIVE ATTITUDE TO X*

A to B Attitude Varied					A to X Attitude Varied					B to X Attitude Varied				
	Situation	Mean Rating	SD	Comp Sig		Situation	Mean Rating	SD	Comp Sig		Situation	Mean Rating	SD	Comp Sig
1 a	1,3,3	40	19	a/c 01	10 a	3,1,3	56	20	a/c 01	19 a	3,3,1	49	17	a/c 01
b	2,3,3	32	18	a/b 01	b	3,2,3	37	17	a/b 01	b	3,3,2	34	17	a/b 05
c	3,3,3	28	18	b/c NS	c	3,3,3	28	18	b/c 01	c	3,3,3	28	18	b/c 01
2 a	1,3,2	47	15	a/c 01	11 a	3,1,2	48	15	a/c 01	20 a	3,2,1	46	13	a/c 01
b	2,3,2	-	-	-	b	3,2,2	39	13	a/b 01	b	3,2,2	39	13	a/b 01
c	3,3,2	54	17	-	c	3,3,2	54	17	b/c 01	c	3,2,3	37	17	b/c NS
3 a	1,3,1	48	13	a/c NS	12 a	3,1,1	47	15	a/c NS	21 a	3,1,1	47	15	a/c 01
b	2,3,1	42	15	a/b 01	b	3,2,1	46	13	a/b NS	b	3,1,2	48	15	a/b NS
c	3,3,1	49	17	b/c 01	c	3,3,1	49	17	b/c NS	c	3,1,3	56	20	b/c 01
4 a	1,2,3	42	19	a/c 01	13 a	2,1,3	54	17	a/c 01	22 a	2,3,1	46	15	a/c 01
b	2,2,3	-	-	-	b	2,2,3	-	-	-	b	2,3,2	-	-	-
c	3,2,3	37	17	-	c	2,3,3	32	18	-	c	2,3,3	32	18	-
5 a	1,2,2	46	13	a/c 01	14 a	2,1,2	52	12	-	23 a	2,2,1	46	15	-
b	2,2,2	43	15	a/b NS	b	2,2,2	43	15	a/b 01	b	2,2,2	43	15	a/b NS
c	3,2,2	39	13	b/c 05	c	2,3,2	-	-	-	c	2,2,3	-	-	-
6 a	1,2,1	45	11	a/c NS	15 a	2,1,1	48	14	a/c NS	24 a	2,1,1	48	14	a/c 01
b	2,2,1	46	16	a/b NS	b	2,2,1	46	16	a/b NS	b	2,1,2	52	12	a/b 05
c	3,2,1	46	13	b/c NS	c	2,3,1	46	15	b/c NS	c	2,1,3	56	17	b/c NS
7 a	1,1,3	43	18	a/c 01	16 a	1,1,3	43	18	a/c NS	25 a	1,3,1	48	13	a/c 01
b	2,1,3	54	17	a/b 01	b	1,2,3	42	19	a/b NS	b	1,3,2	47	15	a/b NS
c	3,1,3	56	20	b/c NS	c	1,3,3	40	19	b/c NS	c	1,3,3	40	19	b/c 01
8 a	1,1,2	45	14	a/c NS	17 a	1,1,2	45	14	a/c NS	26 a	1,2,1	45	11	a/c NS
b	2,1,2	52	12	a/b 01	b	1,2,2	46	13	a/b NS	b	1,2,2	46	13	a/b NS
c	3,1,2	48	15	b/c 05	c	1,3,2	47	15	b/c NS	c	1,2,3	42	19	b/c NS
9 a	1,1,1	46	15	a/c NS	18 a	1,1,1	46	15	a/c NS	27 a	1,1,1	46	15	a/c NS
b	2,1,1	48	14	a/c NS	b	1,2,1	45	11	a/b NS	b	1,1,2	45	14	a/b NS
c	3,1,1	47	15	b/c NS	c	1,3,1	48	13	b/c NS	c	1,1,3	43	18	b/c NS

* The tables contain three main columns: in the first column A's attitude to B is varied systematically, in the second, A's attitude to *X* is varied systematically, and in the third column B's attitude to *X* is varied systematically. All columns have identical entries. a) The first entry identifies the cell and situation within the cell; the cell is identified by a number and the situation by a letter. b) The second entry gives the strength of the attitudes holding for the three dyads constituting the situation in numbers. Two of the possible 27 situations were not presented to subjects; hence no data concerning them were available. c) The third entry gives the mean rating of 100 subjects to the situation. In the experiment under discussion a neutral rating was scored as zero and the degree of pleasantness or unpleasantness was scored by positive or negative numbers away from zero. For purposes of ease of calculation, among others, the zero rating was transformed to 55 and all other ratings were transformed accordingly, the higher the score the more unpleasant the situation. Where no data were available, a dash is entered in the proper place. d) The fourth entry gives the standard deviation of the ratings for each situation. Where no data were available, a dash is entered. e) The fifth entry gives the significance of the difference between the respective means of the three possible paired comparisons between the situations within each cell as 01, 05, or NS (not significant). Where the data are missing for a paired comparison, a dash is entered (6, p. 6).

Note: The data in Table 1 are adapted from "Positive and negative affect as a function of perceived discrepancy in AB*X* situations," by Price, Harburg, and McLeod (6), and are used by permission of *Human Relations* and the senior author.

TABLE 2

THE EFFECTS OF VARIATION OF ATTITUDE INTENSITY ON ABX SITUATIONS OF THE FORM: A HAS A POSITIVE ATTITUDE TO B; A HAS A POSITIVE ATTITUDE TO X; B HAS A NEGATIVE ATTITUDE TO X*

A to B Attitude Varied					A to X Attitude Varied					B to X Attitude Varied				
	Situation	Mean Rating	SD	Comp Sig		Situation	Mean Rating	SD	Comp Sig		Situation	Mean Rating	SD	Comp Sig
1 a	1,3,3	69	20	a/c 01	10 a	3,1,3	56	19	a/c 01	19 a	3,3,1	76	12	a/c 01
b	2,3,3	76	17	a/b 01	b	3,2,3	76	15	a/b 01	b	3,3,2	71	15	a/b 01
c	3,3,3	85	14	b/c 01	c	3,3,3	85	14	b/c 01	c	3,3,3	85	14	b/c 01
2 a	1,3,2	64	18	a/c 01	11 a	3,1,2	56	19	a/c 01	20 a	3,2,1	67	13	a/c 01
b	2,3,2	70	13	a/b 01	b	3,2,2	68	13	a/b 01	b	3,2,2	68	13	a/b NS
c	3,3,2	71	15	b/c NS	c	3,3,2	71	15	b/c NS	c	3,2,3	76	15	b/c 01
3 a	1,3,1	65	14	a/c 01	12 a	3,1,1	53	14	a/c 01	21 a	3,1,1	53	14	a/c NS
b	2,3,1	71	13	a/b 01	b	3,2,1	67	13	a/b 01	b	3,1,2	56	14	a/b NS
c	3,3,1	76	12	a/c 01	c	3,3,1	76	12	b/c 01	c	3,1,3	56	19	b/c NS
4 a	1,2,3	69	17	a/c 01	13 a	2,1,3	-	-	-	22 a	2,3,1	71	13	a/c 05
b	2,2,3	70	18	a/b NS	b	2,2,3	70	18	-	b	2,3,2	70	13	a/b NS
c	3,2,3	76	15	b/c 01	c	2,3,3	76	17	b/c 05	c	2,3,3	76	17	b/c 05
5 a	1,2,2	62	12	a/c 01	14 a	2,1,2	-	-	-	23 a	2,2,1	66	11	a/c NS
b	2,2,2	63	9	a/b NS	b	2,2,2	63	9	-	b	2,2,2	63	9	a/b 05
c	3,2,2	68	13	b/c 01	c	2,3,2	70	13	b/c 01	c	2,2,3	70	18	b/c 01
6 a	1,2,1	63	11	a/c 05	15 a	2,1,1	-	-	-	24 a	2,1,1	-	-	-
b	2,2,1	66	11	a/b NS	b	2,2,1	66	11	-	b	2,1,2	-	-	-
c	3,2,1	67	13	b/c NS	c	2,3,1	71	13	b/c 01	c	2,1,3	-	-	-
7 a	1,1,3	60	16	a/c NS	16 a	1,1,3	60	16	a/c 01	25 a	1,3,1	65	14	a/c NS
b	2,1,3	-	-	-	b	1,2,3	69	17	a/b 01	b	1,3,2	64	18	a/b NS
c	3,1,3	56	19	-	c	1,3,3	69	20	b/c NS	c	1,3,3	69	20	b/c NS
8 a	1,1,2	59	12	a/c NS	17 a	1,1,2	59	12	a/c 05	26 a	1,2,1	63	11	a/c 01
b	2,1,2	-	-	-	b	1,2,2	62	12	a/b NS	b	1,2,2	62	12	a/b NS
c	3,1,2	56	14	-	c	1,3,2	64	18	b/c NS	c	1,2,3	69	17	b/c 01
9 a	1,1,1	59	10	a/c 01	18 a	1,1,1	59	10	a/c 01	27 a	1,1,1	59	10	a/c NS
b	2,1,1	-	-	-	b	1,2,1	63	11	a/b 01	b	1,1,2	59	12	a/b NS
c	3,1,1	53	14	-	c	1,3,1	65	14	b/c NS	c	1,1,3	60	16	b/c NS

* See footnote for Table 1.

Note: The data in Table 2 are adapted from "Positive and negative affect as a function of perceived discrepancy in ABX situations," by Price, Harburg, and McLeod (6), and are used by permission of *Human Relations* and the senior author.

C. TESTING OF HYPOTHESES

Since the basic nature of ABX situations are determined by attitude configurations but not by attitude intensities, it is proper to talk about a basic situation defined by the signs of the attitudes but not by their intensities. Four basic situations were used in this analysis: $A + LB$, $A + LX$, $B + LX$; $A + LB$, $A + LX$, $B - LX$; $A + LB$, $A - LX$, $B + LX$; and $A + LB$, $A - LX$, $B - LX$. But they were not varied identically. There are 25 variations of the first basic situation, 24 variations of the second, six of the third, and only three of the fourth (which add up to 58). Table 1 presents all the published data with respect to the first basic situation, where all the attitudes are positive.

The first column of Table 1 "tests" the predictions of the first hypothesis. Since the situation is basically "pleasant," the degree of rated pleasantness of the situation should be a function of the degree of positive attitude strength between A and B. Note that cells 2 and 4 do not fit this analysis, since they only have two entries. Inspection shows that none of the remaining cells in the column supports the hypothesis. The ratings of the three situations in cells 6 and 9 are not significantly different from each other. In cells 1 and 7 the ratings of the situation with moderate attitude intensity (situation b) are not significantly different from the ratings of the situation with high attitude intensity (situation c). In cell 5 the rating of the situation with moderate attitude intensity (situation c) is not significantly different from the rating of the situation with low attitude intensity (situation a). And finally, in cells 3 and 8 the ratings of the situation with low attitude intensity (situation a) are not significantly different from the ratings of the situation with high attitude intensity (situation c), while the ratings of the situation with moderate attitude intensity are significantly different from both; in one case it is rated as being more pleasant and in the other as less pleasant.

The second and third columns "test" the second hypothesis. If situations with equal attitude intensities are considered as having zero discrepancy, most of the discrepancy of the cells changes range from zero to 2. There are four cells—11, 17, 20, and 26—where the discrepancy values are zero, 1, 1. For these the criterion of intensity is applied (6, p. 98); the situation whose sum of affect strength is greater with the same degree of discrepancy will have a stronger effect upon affect.

Again we find four cells that do not fit the analysis for lack of data—13, 14, 22, 23. Inspection shows that only one of the remaining 14 cells supports the second hypothesis; the other 13 cells do not support the hypothesis. Cell 10 is the confirming cell; the ratings of the situations in this cell are all significantly different and the differences are in the predicted directions.

In seven cells—12, 15, 16, 17, 18, 26, 27—the differences in the ratings of their constituent situations are not statistically significant. In cell 21 the rating of the situation with zero discrepancy is not significantly different from the rating of the situation with a discrepancy of 1. In cells 24 and 25 the ratings of the situation with a discrepancy of 1 are not significantly different from the ratings of the situation with a discrepancy of 2. In cell 19 the situation with a discrepancy of 2 is significantly preferred to a situation with a discrepancy of 1. In cells 11 and 20 the discrepancy distribution of zero, 1, 1 is found. The ratings of both situations with a discrepancy of 1 are not significantly different in cell 20; and the situation with attitude strength distribution 1, 2 (situation a) is significantly preferred over the situation with attitude strength distribution 3, 2 (situation c). In both these cases the criterion of intensity would predict that the situation with an attitude intensity distribution 3, 2 would be preferred to a situation with a distribution of 1, 2.

To summarize the findings of Table 1: Of the 21 sets of triads "testing" the two hypotheses, two sets support the hypotheses, whereas 19 sets do not.

Table 2 presents the data with respect to the second basic situation where B has a negative attitude towards *X* while A has a positive attitude (assuming A likes B). This is an imbalanced situation that is basically experienced as unpleasant; hence, the hypotheses predict the opposite effect upon the ratings than in the preceding basic situation: e.g., the ratings will fall to indicate lesser unpleasantness rather than rise to indicate lesser pleasantness.

Again, the first column "tests" the first hypothesis. Inspection shows that three cells—7, 8, 9—are not fit for analysis because of lack of data. Of the remaining six cells, two—1, 3—confirm the hypothesis: the differences in ratings are significant in the proper directions; three—2, 4, 5—disconfirm the hypothesis, and one—6—is ambiguous. In cells 4 and 5 the rating of the situation with attitude intensity 2 is not significantly different from the rating of the situation with attitude intensity 1. In cell 2 the rating of the situation with attitude intensity 2 is not significantly different from the rating of the situation with attitude intensity 3. Cell 6 is ambiguous, since the rating of the situation with attitude intensity 2 is roughly midway between the ratings of the other two situations, but it is not significantly different from either of them.

The remaining two columns "test" the second hypothesis. Four cells—13, 14, 15, 24—are not fit for analysis because of lack of data. Two cells—10, 12—support the hypothesis. Twelve cells—11, 16, 17, 18, 19, 20, 21, 22, 23, 25, 26, 27—do not. The differences between the ratings of the situations in three cells—21, 25, 27—are not significant. In cells 11, 16, 17, 20, and 26 the ratings of the situation with a discrepancy of 1 (c) are not significantly different from the ratings of the situation with a discrepancy of zero (b). In cells 18 and 22 the ratings of the situation with a discrepancy of 1 are not significantly different from the ratings of the situation with a discrepancy of 2.

Cell 19 is an interesting case. Although the ratings of the three situations there are all significantly different from each other, the effect of a discrepancy of 1 is far stronger than the effect of a discrepancy of 2, which is not in accord with the hypothesis. Cell 23 is another instance of a distribution of values zero, 1, 1 where the criterion of intensity predicts a difference not found in the data.

To summarize the findings of Table 2: Of the 20 sets of triads "testing" the two hypotheses, four support the hypotheses, 15 do not support the hypotheses, and one set is ambiguous.

Of the six situations of the basic form—A likes B, A dislikes *X*, B likes *X*—presented to the subjects it is possible to construct only two groups of three situations to test the effect of change in attitude intensity. They are (*a*) 3, 3, 1 with a mean rating of 63, and standard deviation of 14; (*b*) 3, 3, 2 with a mean rating of 69, and standard deviation of 15; (*c*) 3, 3, 3 with a mean rating of 72, and standard deviation of 19; and (*d*) 1, 3, 1 with a mean rating of 58, and standard deviation of 13; (*e*) 1, 3, 2 with a mean rating of 59, and standard deviation of 15; (*f*) 1, 3, 3 with a mean rating of 60, and standard deviation of 18. In the first group the rating of (*a*) is significantly different from that of (*b*) and (*c*), but (*b*) is not significantly different from that of (*c*). In the second group the ratings are not significantly different. Both these tests disconfirm the hypotheses.

It is not possible to construct a proper triad of situations from those submitted to the subjects in the fourth basic form: A likes B and both dislike *X*.

The data presented can now be summarized. Of the 43 triple comparisons analyzed, 36 disconfirm the hypotheses, one neither confirms nor disconfirms the hypotheses, and only six confirm the hypotheses.

The evidence is clear-cut. Had the investigators studied all the data available to them rather than an arbitrarily selected part of the data they would have had summarily to reject their hypotheses. As the hypotheses stand, they are untenable.

D. Discussion

Two substantive lessons seem to be gleanable from this incident. The first lesson concerns the "passing off" of hypotheses as being confirmed, at least in part, when a more careful analysis of the available data clearly shows that this is not the case. It is too easy and factually incorrect to attribute this merely to the carelessness of the individual investigators who reported the above research. The paper criticized was based upon a doctoral dissertation, which means the data and their analysis were under the supervision of a senior professor and were subjected to a critical review by an examining committee where the research was presumably defended by the candidate for the Ph.D. degree. In addition, it was afterwards accepted for publication by a professional journal and presumably submitted to a critical review by the editorial staff of that journal and its readers. Clearly many professional psychologists were involved in the process whereby this paper was printed and all should assume some degree of responsibility for its publication.

It seems to me that this publication is a symptom of an unhealthy tendency that is currently quite prevalent in the profession, a tendency that seems to assume that it is somewhat shameful or "bad" to formulate hypotheses to be *disconfirmed* by subsequent research. In fact, judging from this paper, it seems to hold that even graduate students are "forced" to formulate "true" hypotheses to be confirmed by subsequent research in order to be granted a degree. One would expect, however, that the probability of relatively inexperienced graduate students formulating new "true" hypotheses should be low. We seem to have forgotten that research that clearly disconfirms a scientific hypothesis is as valuable for science as the confirmation of hypotheses. What is of primary and sole importance is the carefully thought through rationale for formulating the hypothesis submitted for investigation and the careful research method employed to test the degree by which, if any, the hypothesis merits further consideration.

Similar difficulties seem to have troubled the better established physical sciences in their earlier days. Popper (5, p. 82) cites from a chemistry textbook by J. Black, published in 1803: "A nice adaptation of conditions will make almost any hypothesis agree with the phenomena. This will please the imagination but does not advance our knowledge."

The second lesson also points to undesirable consequences that stem from another tendency that prevails in contemporary scientific psychology, the tendency that assumes that we cannot truly become a science unless and until we can so quantify psychological phenomena that they become amenable to mathematical analysis. And not any kind of mathematical analysis will do, but primarily the kind that leads to continuous curves. Hence, we formulate our hypotheses in the general form: The more there is of x (the independent variable) the more—or less, as the case may be—there will be of y (the dependent variable). We become intrinsically unable to cope with discontinuities and consequently gloss over them.

The data from this experiment, which conform to data from other experiments too numerous to mention here, do show that there are probably many areas in psychology where simple continuity is not the case. It is true that for many areas in psychology the formulation: *the more there is of x the more, or less, there will be of y,* holds; it is just as true that it does not hold for many other areas in psychology, and cognitive organization is obviously one of them. But how can this ever be found out if data are so analyzed to stress the seeming continuity and obscure the discontinuities? And, unless we learn to discipline ourselves to obtain all that we can from data that is always laboriously gathered, we are doomed to prejudge facts rather than let the facts lead us to correct, true judgments.

REFERENCES

1. BROWN, R. Social Psychology. New York: Free Press, 1965.
2. NEWCOMB, T. M. An approach to the study of communicative acts. *Psychol. Rev.*, 1953, **60**, 393-404.
3. ———. The cognition of persons as cognizers. In R. Tagiuri & L. Petrullo (Eds.), *Person Perception and Interpersonal Behavior*. Stanford, Calif.: Stanford Univ. Press, 1958.
4. ———. Individual systems of orientation. In S. Koch (Ed.), *Psychology: A Study of a Science* (*Vol. 3*). New York: McGraw-Hill, 1959.
5. POPPER, K. R. The Logic of Scientific Discovery. New York: Basic Books, 1959.
6. PRICE, K. O., HARBURG, E., & MCLEOD, J. M. Positive and negative affect as a function of perceived discrepancy in AB*X* situations. *Hum. Relat.*, 1965, **18**, 87-100.

Economic and Political Studies Division
Institute for Defense Analyses
400 Army-Navy Drive
Arlington, Virginia 22202